University of Alberta

Artificial Neural Network Modelling of Filtration Performance

by

Ronald-Ray Tiñana Tupas ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment
of the requirements for the degree of **Master of Science**

in

Environmental Science

Department of Civil and Environmental Engineering

Edmonton, Alberta

Fall 2000

National Library of Canada

Acquisitions and Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada

Acquisitions et services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-59890-X

Canada

University of Alberta

Library Release Form

Name of Author: Ronald-Ray Tiñana Tupas

Title of Thesis: Artificial Neural Network Modelling of Filtration Performance

Degree: Master of Science

Year this Degree Granted: 2000

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

12835-103 St.
Edmonton, AB
T5E-4L9

July 13, 2000

"Follow your dreams. You can reach your goals. I'm living proof. Beefcake! BEEFCAKE!"

- Eric Cartman

Abstract

In water treatment, there is a constant trend of ever increasingly stringent standards for finished water quality and with that comes the need to meet and surpass these standards, which include particle count levels. This requires a good knowledge and understanding of the treatment processes involved as well as good control of these processes. However, finished water particle counts are highly variable and difficult to control. In order to improve process control, an artificial neural network (ANN) model of filtration performance was developed for the E.L. Smith Water Treatment Plant (WTP) in Edmonton, Alberta. This model is unique because it encompasses the complete treatment plant rather than focussing on a single unit treatment process. The model was also applied and demonstrated to be a potentially powerful tool in assisting WTP operators in optimising process control and researchers in analysing particle counts in a virtual lab setting.

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled *Artificial Neural Network Modelling of Filtration Performance* submitted by *Ronald-Ray Tiñana Tupas* in partial fulfilment of the requirements for the degree of *Master of Science* in *Environmental Science*.

Dr. Stephen Stanley

Ian Buchanan

Dr. Ian Buchanan

John Feddes

Dr. John Feddes

June 7, 2000

Acknowledgements

I would like to express my sincere thanks to a number of people who helped make completion of this thesis possible:

- My supervisor Dr. Stephen Stanley for his excellent guidance and enthusiastic support for my project.
- The American Water Works Association Research Foundation and EPCOR Water Services Inc. for their financial support.
- Riyaz Shariff, Simon Thomas, and Tina Mendes from EPCOR Water Services for their assistance in obtaining data and answering countless questions about the treatment plant.
- Qing Zhang from the Department of Civil and Environmental Engineering at the University of Alberta for his help on a lot of the technical aspects of artificial neural networks.
- All those that I've met and befriended through my Master's program for helping me survive through this journey.
- And last but not least, my mom and dad for their incredible love and support throughout my entire university career. Correction, my entire life!

Table of Contents

1.0 Introduction **1**
  1.1 Background 1
  1.2 Objectives 4
  1.3 Scope of Study 5

2.0 Background **6**
  2.1 Introduction 6
  2.2 Filtration Performance Parameters 7
    2.2.1 Descriptions 7
    2.2.2 Significance of Filtration Performance Parameters 8
  2.3 Filtration Process and Mechanisms 13
    2.3.1 Description of Filtration 13
    2.3.2 Mechanisms of Filtration 14
    2.3.3 Filter Operation and Control 17
  2.4 Factors Affecting Filtration Performance 19
    2.4.1 Operational Parameters 19
      2.4.1.1 Filtration Rate 19
      2.4.1.2 Filter Media Size 20
      2.4.1.3 Filter Media Depth 21
    2.4.2 Solution Chemistry 21
      2.4.2.1 Coagulant Dosage and Type 21
      2.4.2.2 Polymer Dosage and Type 23
      2.4.2.3 pH and Alkalinity 24
      2.4.2.4 Ionic Strength 28
      2.4.2.5 Particle Concentration 29
      2.4.2.6 Hardness and Water Softening 30
      2.4.2.7 Organic Matter Concentration 31
      2.4.2.8 Activated Carbon Dosage 33
      2.4.2.9 Temperature 33
    2.4.3 Suspended Particle Properties 34
      2.4.3.1 Zeta Potential 34
      2.4.3.2 Particle Size and Size Distribution 36
  2.5 Aspects of Artificial Neural Networks 38
    2.5.1 Description 38
    2.5.2 Architecture, Components, and Aspects of ANN 40
      2.5.2.1 Neurons 40
      2.5.2.2 State of Activation 42
      2.5.2.3 Set of Connection Weights 42
      2.5.2.4 Propagation Function 42
      2.5.2.5 Activation Function 43
      2.5.2.6 Learning Rules and Network Learning 44
    2.5.3 Generalisation and Convergence 45
    2.5.4 Kohonen Categorisation 46
  2.6 Other AI Systems 48

- 2.6.1 Expert Systems 49
- 2.6.2 Fuzzy Logic 50
- 2.6.3 Hybrid Systems 51

2.7 E.L. Smith WTP Description 53

3.0 Methodology **56**

3.1 Introduction 56

3.2 Source Data Analysis 58

- 3.2.1 Data Availability 58
- 3.2.2 Input and Output Selection 58
- 3.2.3 Time Scale Analysis 62
- 3.2.4 Data Characteristics 64

3.3 Methodology of Data Categorisation and Preparation 66

- 3.3.1 First Stage Categorisation 66
- 3.3.2 Second Stage Categorisation 70
  - *3.3.2.1 Category 1* 70
  - *3.3.2.2 Category 2* 73
- 3.3.3 Noise Analysis 75
- 3.3.4 Data Set Extraction Preparation 90

3.4 Model Development and Optimisation 93

- 3.4.1 Architecture 93
- 3.4.2 Modelling Parameters 95
- 3.4.3 Model Amalgamation 101

4.0 Results **105**

4.1 Model Performance Results 105

4.2 Model Predictions 112

- 4.2.1 Model Predictions With One Varied Input 113
- 4.2.2 Model Predictions With Two Varied Inputs 118

5.0 Discussion **125**

5.1 Performance Summary 125

5.2 Model Applications 127

5.3 Conclusions 130

5.4 Recommendations 133

References **136**

Appendix **141**

List of Tables

Table 3-1. Available data at E.L. Smith WTP..........59

Table 3-2. Statistics of input and output data to be used..........65

Table 3-3. Statistical values of input data in each category..........67

Table 3-4. Values of extreme cases compared to category average..........69

Table 3-5. Statistical values of input data in each sub-category of category 1..........71

Table 3-6. Values of extreme cases compared to sub-category average..........72

Table 3-7. Statistical values of input data in each sub-category of category 2..........73

Table 3-8. Values of extreme cases compared to sub-category average..........75

Table 3-9. Number of noise patterns removed from the category data sets..........76

Table 3-10. Model performance results testing various network architectures..........94

Table 3-11. Model performance results testing various combinations of numbers of hidden neurons..........95

Table 3-12. Best model performance results from the testing of various combinations of activation functions..........96

Table 3-13. Best model performance result after testing various combinations of learning rate, momentum, and initial weights..........99

Table 3-14. Model performance results of testing various scale functions..........100

Table 3-15. Model performance results of the category 1 sub-category models compared to the combined model..........102

Table 3-16. Model performance results of the category 2 sub-category models compared to the combined model..........102

Table 3-17. C1S1 model performance results with the inclusion of noise patterns and extreme case patterns..........103

Table 3-18. C1S2 model performance results with the inclusion of noise patterns and extreme case patterns..........103

Table 3-19. C2 model performance results with the inclusion of noise patterns and extreme case patterns..............104

Table 4-1. Spring/fall (C1S1) data set characteristics..............106

Table 4-2. Spring/fall (C1S1) model performance results..............106

Table 4-3. Winter (C1S2) data set characteristics..............108

Table 4-4. Winter (C1S2) model performance results..............108

Table 4-5. Summer (C2) data set characteristics..............110

Table 4-6. Summer (C2) model performance results..............111

Table 4-7. Values of inputs held constant during examination of effects of inputs on particle counts..............113

Table 4-8. Overview of the impact of parameters on finished water particle counts....118

Table 5-1. Statistical performance of the models that were developed..............125

List of Figures

Figure 2-1. Alum coagulation diagram and zeta potential with charge neutralisation zones........ 25

Figure 2-2. Optimal water treatment configuration as a function of raw water characteristics........29

Figure 2-3. Electrical double layer of a negatively charged colloid........35

Figure 2-4. A basic artificial neural network........38

Figure 2-5. A basic neuron........41

Figure 2-6. Process schematic of E.L. Smith WTP........55

Figure 3-1. Minute readings vs. hourly average filter effluent particle counts from April 15-17, 1998 for filter 5 at E.L. Smith WTP........ 64

Figure 3-2. Hourly average vs. daily average reservoir influent particle counts from E.L. Smith WTP........64

Figure 3-3. Distinction between category 1 and 2 based on turbidity and temperature...67

Figure 3-4. Distinction between category 1 and 2 based on turbidity and temperature after manual placement........ 68

Figure 3-5. Comparison of extreme data points to category 1 and 2 based on alkalinity and hardness........69

Figure 3-6. Comparison of extreme values to sub-category 1 and 2 of category 1 based on PAC dose and alum dose........ 72

Figure 3-7. Comparison of extreme values to sub-category 1 and 2 of category 2 based on lime dose and alum dose........ 74

Figure 3-8. Actual vs. model predicted particle counts using a Gaussian and tanh 15 activation function and production set data........ 97

Figure 3-9. Actual vs. model predicted particle counts using a Gaussian and tanh 15 activation function and the entire cat 1, sub-cat 1data set........ 97

Figure 3-10. Actual vs. model predicted particle counts using the logistic activation function and production set data........ 98

Figure 3-11. Actual vs. model predicted particle counts using the logistic activation function and the entire cat 1, sub-cat 1data set 98

Figure 4-1. Spring/fall (C1S1) model predictions vs. actual particle count data 107

Figure 4-2. Winter (C1S2) model predictions vs. actual particle count data 109

Figure 4-3. Summer (C2) model predictions vs. actual particle count data 111

Figure 4-4. The effect of alum dose on particle counts as predicted by each model 114

Figure 4-5. The effect of lime dose on particle counts as predicted by each model 115

Figure 4-6. The effect of PAC dose on particle counts as predicted by each model 116

Figure 4-7. The effect of polymer dose on particle counts as predicted by each model 117

Figure 4-8. The effect of alum and lime dose on particle counts as predicted by the spring/fall model 119

Figure 4-9. The effect of alum and lime dose on particle counts as predicted by the winter model 120

Figure 4-10. The effect of alum and lime dose on particle counts as predicted by the summer model 120

Figure 4-11. The effect of alum and polymer dose on particle counts as predicted by the spring/fall model 121

Figure 4-12. The effect of alum and polymer dose on particle counts as predicted by the winter model 122

Figure 4-13. The effect of alum and polymer dose on particle counts as predicted by the summer model 123

Figure 5-1. Overall configuration of the predictive model 134

Figure 5-2. Configuration of predictive model with all ANN models combined into one 134

List of Abbreviations and Symbols

@	at
%	percent
°C	degree Celsius
AI	artificial intelligence
alk	alkalinity
ANN	artificial neural network
bp	backpropagation
btw	between
C1, C2, C3	clarifier 1, clarifier 2, clarifier 3
C1S1	category 1, sub-category 1
C1S2	category 1, sub-category 2
C2S1	category 2, sub-category 1
C2S2	category 2, sub-category 2
C2 model	category 2 model
counts/mL	particle counts per millilitre
diam	diameter
EBC	European Brewery Convention
eqn	equation
ESWTR	Enhanced Surface Water Treatment Rule
h	hour
h/d	hour per day
hard	hardness

Hz	hertz
iep	isoelectric point
init wt, ini wt	initial weight
l rate	learning rate
M	molar or mole per litre
m/h	metre per hour
med	medium
mg/L	milligram per litre
mL	millilitre
ML/d	megalitre per day
mm	millimetre
mM	millimolar or millimole per litre
mV	millivolt
NTU	nephelometric turbidity units
outln	outlined
OD	outer diameter
PAC	powdered activated carbon
psi	pounds per square inch
R^2	coefficient of multiple determination
recarb effl	recarbonation effluent
res inf	reservoir influent
RH	relative humidity
s	second

std dev	standard deviation
temp	temperature
turb	turbidity
µm	micrometre
wt	weight
WTP	water treatment plant

1.0 Introduction

1.1 Background

One of the most important goals of a water treatment plant (WTP) is the removal of particulate matter. *Giardia* and *Cryptosporidium* removal are of particular concern in water treatment, and this is addressed by stricter standards such as the Enhanced Surface Water Treatment Rule (ESWTR) (Lind 1997), which may ultimately include standards for particle count levels in finished water. With this trend of ever increasingly stringent standards for finished water quality comes the need to meet and surpass these standards. This requires a good knowledge and understanding of the treatment processes involved as well as good control of these processes. Filtration in particular is important in that it removes particles from the water and is credited with the removal of 99% of *Giardia* cysts present in the water. While much study has been done to understand the theory behind the process as well as the factors involved in influencing the performance of the process, there is still uncertainty in terms of understanding the quantitative relationships between various factors and filtration performance. More specifically, the relationship between various factors and particle counts, which is a direct measure of filtration performance, is still not completely understood. This is due to the fact that such relationships are quite complex. In fact, there are situations in which the reason for observed particle count behaviour is unknown (Ginn Jr., Bennett and Wheatley 1997).

In practice, an understanding of how these various factors influence filtration performance is important since this will lead to better control of the treatment process. Currently, process control stems from the results of traditional jar tests performed to determine chemical dosages as well as from operator experience. However, jar tests only assess clarification performance and does not directly assess filter performance. While this generally works for controlling turbidity, particle counts are more sensitive to small changes in filtration efficiency than turbidity (LeChevallier and Norton 1992; Hall and Croll 1997; Lind 1997), and thus particle counts are more variable and are harder to control. Therefore, with better control and improved understanding of these relationships, stringent water quality standards can be met.

One way of understanding these relationships is through deriving mathematical models and equations. However, determining an exact mathematical model to describe the relationships between various factors and filtration performance is very difficult if not impossible given that such relationships are complex and non-linear. Therefore, a different type or method of modelling is necessary rather than conventional mathematical modelling. Artificial neural network (ANN) modelling is just such an alternative. It is applicable to problems in which the cause-effect relationships are complex and non-linear and where no mathematical formulae exist, such as the case with filtration performance. If enough data are available that represent all aspects of the problem domain, then a model can successfully be developed. Unfortunately, ANN will not give an explanation of why different effects are observed. However, it is up to the investigator to interpret these observations. Therefore, it is important to have expert knowledge and a strong understanding of the problem domain, and in this case, a strong understanding of the filtration process itself not only to interpret results, but also to develop a model that accurately represents all aspects involved in filtration. ANN also will not provide a precise explanation as to how the relationships are derived, but that is not necessary for the purposes of this study since determining the relationships and interpreting them are the more important goals in this study.

Development of an ANN model for filtration performance addresses the problems of meeting stringent water quality standards by providing a description of the complex relationships involved in filtration performance and improving the process control of water treatment. Once developed, the optimised model can be used to provide insight into particle count behaviour without experimentation in a standard wet laboratory or through pilot tests. Control of the treatment process would improve since the ANN model could provide immediate multiple treatment options as opposed to performing time consuming jar tests. The model can also be implemented online with the process control system in an advanced control mode improving process control by not only providing treatment options but by initiating the action. Improved control results in the optimisation of the treatment process and will lead to improved water quality. Moreover, there are savings benefits from reagent costs since an optimal chemical dose can be determined without utilising an excess amount of reagents. In addition to the quickness and convenience of using the ANN model for process control, the model could also be useful in operator training by simulating possible scenarios in which the operator would learn the results of various treatment options. All of these various uses show the tremendous benefits of developing and utilising an ANN model of filtration performance.

1.2 Objectives

One of the main objectives in this study is to determine the feasibility of developing an ANN model of filtration performance. Furthermore, the feasibility of developing a model of not just a single unit process but for a complete plant will be determined since filtration performance will be modelled based on plant operating conditions and raw water quality. This is significant because even though research has been conducted on the ANN modelling of unit processes in water treatment (Mirsepassi, Cathers and Dharmappa 1995; Maier and Dandy 1996; Rodriguez and Serodes 1996; Gagnon, Grandjean and Thibault 1997; Han et al. 1997; Rodriquez, Serodes and Cote 1997; Stanley and Zhang 1997; Zhang and Stanley 1997), little development has been done in developing process models for a complete WTP. Therefore, the model to be developed in this study is unique with beneficial implications as a result of this type of model development.

Other objectives of this study are:

- to demonstrate the use of the ANN model as a tool for WTP operators and researchers
- to examine the effects of various parameters on particle counts

1.3 Scope of Study

To meet the objectives, the following work was completed. A literature review of filtration and the factors that affect the process was conducted and the available data were assessed as part of the model development process. From there, models were developed and optimised to predict with statistically reasonable results for the E.L. Smith WTP in Edmonton, Alberta. Finally, the applicability of the model and the benefits of using the model were demonstrated. Particle count behaviour in response to changing various factors was examined by utilising the developed model.

2.0 Background

2.1 Introduction

Before development of an ANN model can begin, background information of the problem domain needs to be researched and examined in order to have an understanding of the problem domain and subsequently to be able to develop a representative model. As a result, presented in this section is an overview of filtration performance and the filtration process as well as a discussion of the ANN modelling technique.

The first topic to be discussed is filtration performance parameters. An understanding of how performance is measured and their significance is important since that is what is being modelled. Next, the filtration process and the mechanisms involved are discussed followed by a discussion on the factors that affect filtration performance. This is necessary to be able to determine the parameters that are to be used in model development. In order to be able to design, develop and optimise an ANN model, an understanding of ANN and its aspects are necessary, and that is covered in the following section. Other artificial intelligence (AI) systems are discussed afterward as options to ANN. Finally, the treatment process at E.L. Smith WTP is described since the model is to be specifically developed for that plant.

2.2 Filtration Performance Parameters

2.2.1 Descriptions

Two of the most commonly used performance parameters for filtration are turbidity and more recently particle counts. Turbidity is a light-transmitting property of water caused by a wide variety of suspended materials, which interfere with the passage of light through water (Sawyer, McCarty and Parkin 1994). It is measured in nephelometric turbidity units (NTU), which is a measure of the degree of light scattering from particles. Therefore, it is an indirect measure of the amount of particulate matter that is present in water. The suspended particles responsible for turbidity can be inorganic colloidal rock, silt, or clay materials, or organic materials of either natural or anthropogenic origin. Particle counts, on the other hand, are a direct measure of the amount of particulate matter present in solution that is causing turbidity. Turbidity is usually measured using a turbidimeter. A light source is shined through the water sample in which a photoelectric detector detects the intensity of scattered light at right angles to the path of the incident light (Sawyer, McCarty and Parkin 1994). Particle counters measure particle counts slightly differently. The water sample flows through a small channel in which light, typically a laser beam, shines through. A diode opposite the laser detects the light that is not scattered or obscured. The particle is thus counted and sized depending if the particle is greater than the sensor size detection limit. At E.L. Smith WTP, the turbidimeter used is a Hach Model 2100N and the particle counters are the Met One, PCX model. A more detailed description of how they operate is discussed in the appendix on pages 141-144.

2.2.2 Significance of Filtration Performance Parameters

There are two main reasons for wanting to remove turbidity. One is for aesthetic reasons so that the treated water appears to be clear, not murky. The other reason is for disinfection. Many harmful microorganisms, such as *Giardia* cysts and *Cryptosporidium* oocysts, tend to be adsorbed to particles or encased in solid material, and so it is necessary to remove such particles while disinfectant is used to kill the remaining microbes. Reduction of particle counts is desired for the same reasons as removal of turbidity. Removal of *Cryptosporidium* oocysts is a particularly major concern right now for water treatment. Water disinfectants such as chlorine have little effect on the inactivation of *Cryptosporidium* oocysts, thus filtration is the most practical treatment technology to remove *Cryptosporidium*. However, there is currently no accurate and precise method for determining *Cryptosporidium* removal rates in filtration systems (Li et al. 1997). One study of 66 surface water treatment plants showed that there was no significant correlation between parasite removal and removal of turbidity (LeChevallier, Norton and Lee 1991). Furthermore, it was found in a series of four pilot studies that turbidity removal did not directly correlate well with removal of either *Cryptosporidium* or *Giardia* (Patania et al. 1995). In these pilot studies, it was observed that turbidity was removed to a much lesser extent than *Cryptosporidium* or *Giardia* when raw water turbidity was less than 10 NTU. However, for raw water with higher turbidity, there would be greater log removal of turbidity, thus turbidity is not a proper surrogate for *Cryptosporidium* or *Giardia* (Patania et al. 1995). Nevertheless, the level of turbidity removal did influence the level of organism removal (Patania et al. 1995). Moreover, turbidity was deemed to be a useful predictor of *Giardia* and *Cryptosporidium* removal within a single treatment facility (LeChevallier and Norton 1992). Nonetheless, the U.S. Environmental Protection Agency has proposed stricter particle control standards to expand its Surface Water Treatment Rule to address removal of *Giardia*, *Cryptosporidium*, and other microorganisms by reducing particle counts (Lind 1997). Since *Cryptosporidium* oocysts are roughly spherical particles 4-6 μm in diameter and *Giardia* cysts are approximately 10-20 μm in diameter, it is desirable to reduce particle counts in that size range. Particle counts, however, do not necessarily correlate with turbidity.

Particle counts have been found to be more sensitive to small changes in filtration efficiency than turbidity (LeChevallier and Norton 1992; Hall and Croll 1997; Lind 1997), and removal of particles >5 μm has been considered to be a useful predictor of *Giardia* and *Cryptosporidium* removal (LeChevallier and Norton 1992). For example, in a pilot plant study (Hall and Croll 1997), it was observed that in the first hour of a filtration run where turbidity ranged from 0.15 to 0.2 NTU, the average 2-5 μm particle count ranged from 1500 to 7100 particles/mL. Thus, although there is a slight relationship between turbidity and particle counts, the correlation does not appear to be significant. This is also indicated by particle monitoring that was done at two full-scale plants (Hall and Croll 1997). It was found in this study that although turbidity ranged from 0.1 to 0.2 NTU at both sites, the 2-5 μm size range of particle counts ranged from 20 to 250 particles/mL at one site and 2000 to 6000 particles/mL at the other site. In addition, at the site in which particle counts peaked to 6000 particles/mL from 2000 particles/mL, such changes were not evident from turbidity as it remained fairly constant.

This further indicates the lack of correlation between particle counts and turbidity. In addition, this finding also highlights the greater sensitivity of particle counts with regards to filtration performance compared to turbidity.

Despite the sensitivity of particle counts to changes in filtration performance, the accuracy and interpretation of such data are still not completely certain, particularly with the correlation between particle count removal and *Cryptosporidium* and *Giardia* removal. A study utilising field-scale bag filtration systems declared that removal of 4-6 µm polystyrene microspheres was an accurate and precise surrogate for determining *Cryptosporidium* log removal in bag filtration processes without addition of chemical coagulant (Li et al. 1997). However, naturally occurring 4-6 µm particle counts, as well as 1-25 µm particle counts and turbidity, were found to be less accurate and precise due to great variations in log removals. Variations in particle counts, not just within the bag filtration study, but between treatment plants, can be due to natural variation in water chemistry (Li et al. 1997) and the addition of particles such as floc particles or granular activated carbon fines from other treatment processes (Hall and Croll 1997). Therefore, there is still concern with how reliable a surrogate particle count removal is with *Giardia* and *Cryptosporidium* removal. Although Li *et al.* (1997) found 4-6 µm polystyrene sphere removal to be an accurate surrogate for *Cryptosporidium* removal for bag filtration processes, which was observed under the condition of no added chemical coagulant, pilot studies done by Patania *et al.* (1995) had different results. Direct correlation was not observed between removal of 1-2, 2-5, 5-15, and 1-25 µm particles and *Giardia* cyst or *Cryptosporidium* oocyst removal. Instead, it was observed that removal of particles underestimated removal of *Giardia* and *Cryptosporidium*. However, just as with turbidity, the level of removal of particles can be dependent on concentration in raw water (Patania et al. 1995).

These research studies have illustrated the need for further study to understand the relationship between turbidity and particle counts and how the removal of each quantitatively relates to *Giardia* cyst and *Cryptosporidium* oocyst removal. Even though exact relationships are not clear, general conclusions can still be made. It has been concluded that achieving a turbidity of 0.1 NTU is effective for cyst and oocyst removal (Patania et al. 1995). In addition, it has also been concluded that removal of particles >5 µm has been considered to be a useful predictor of *Giardia* and *Cryptosporidium* removal (LeChevallier and Norton 1992). Therefore, both turbidity and particle counts still remain as important measures of filtration performance.

As stated previously, particle counts are more sensitive to small changes in filtration efficiency than turbidity (LeChevallier and Norton 1992; Hall and Croll 1997; Lind 1997). Particle counts are thus a useful measure within a single treatment plant more so than turbidity because of its sensitivity to these changes. However, with greater sensitivity comes greater variability, and this makes particle counts more difficult to control. An example of this is with the Quarles No. 1 WTP in the state of Georgia (Ginn Jr., Bennett and Wheatley 1997). Particle count peaks were observed with no change in the treatment process during the summer afternoons whereas turbidity remained relatively constant. No certain explanations could be given for this phenomenon. This serves to illustrate the difficulty of controlling particle count levels.. Subsequently, meeting the stricter standards that address *Giardia* and *Cryptosporidium* removal is made difficult. The development of an ANN model for filtration performance for predicting particle counts addresses this concern by improving the process control of water treatment, which subsequently improves the control of particle count levels in finished water.

2.3 Filtration Process and Mechanisms

In addition to understanding the parameters that measure filtration performance, it is necessary to understand how the process works to help determine the factors that have an effect on filtration performance. Consequently, an ANN model representative of the problem domain can be developed. Control of the treatment process is also discussed illustrating the need for improved control to improve performance.

2.3.1 Description of Filtration

Filtration is a unit treatment process used in producing potable water. Its primary objective is to remove particulate matter from the water whether the particles come from the source water or are generated in treatment processes (AWWA 1990). As mentioned previously, filtration performance is typically measured in two ways: filtrate turbidity levels and filtrate particle count levels. Performance can also be measured by headloss through a filter.

There are a number of different types of filters used in potable water production, and consequently they can be generally classified in four ways (AWWA 1990). One way is based on the type of media that is used. Granular media of sand, crushed quartz, anthracite, and garnet can be used in mono-media, dual media, or tri-media form. Instead of a granular bed, a contrasting type of filter is the pre-coat filter where a thin layer of very fine medium such as diatomaceous earth is used. Hydraulic arrangement is another means of classifying filters. With gravity filters, flow through the media achieved by gravity and the filter is open to the atmosphere. On the other hand, pressure filters have the filter medium contained under pressure, and water is driven through the filter by pressure. A third means of classifying filters, particularly granular-bed filters, is by the rate of filtration. Rapid sand filters operate at high flow rates, whereas slow sand filters operate at much lower rates. A final means of classifying filters is by the type of filtration defining where in the filter the solids are removed. If particles are removed within granular material, then this is defined as depth filtration. In contrast, cake filtration is when particles are removed on the entering surface of the filter media as the case with pre-coat filters.

The type of filters used at the E.L. Smith WTP is dual-media granular filters operated under gravity. Anthracite and sand are the two types of media used in the filters. The flow rate is high therefore utilising rapid sand filtration, and since particles are removed within the granular media bed, the type of filtration can be classified as depth filtration as opposed to cake filtration.

2.3.2 Mechanisms of Filtration

The mechanism of rapid filtration is complex in that there are actually multiple mechanisms involved. Particles that are larger than the pore spaces in the filter media are physically removed by the process of interstitial straining (AWWA 1990). However, for smaller particles, filtration is considered to be a combination of particle transport to the media and attachment to the media (O'Melia and Stumm 1967). Small particles must first be brought from the bulk of the fluid within the interstitial spaces of the media to the surface of the media grains by transport mechanisms. Once particles are brought to the surface of media grain, or surface of other deposited solids, physical and chemical forces control the attachment of the particles to the grain surface.

There are three transport mechanisms that are generally considered to be the most important in particle transport: sedimentation, interception, and diffusion (AWWA 1990). Efficiency of particle transport by sedimentation is governed by equation 1:

$$\eta_s = \frac{(\rho_s - \rho) g d_p^2}{18 \mu V} \tag{1}$$

where,

η_s = sedimentation transport efficiency of collector,

ρ_s = filter grain mass density,

ρ = mass density of fluid,

g = acceleration of gravity,

d_p = particle diameter,

μ = fluid viscosity,

V = filtration rate.

Efficiency of particle transport by interception is given by equation 2:

$$\eta_I = 3/2 \left(\frac{d_p}{d_c} \right)^2 \tag{2}$$

where,

η_I = interception transport efficiency of collector,

d_p = particle diameter

d_c = filter grain diameter.

Finally, efficiency of particle transport by diffusion is governed by equation 3:

$$\eta_D = 0.9\left(\frac{KT}{\mu d_p d_c V}\right)^{2/3} \qquad (3)$$

where,

η_D = diffusion transport efficiency of collector,
K = Boltzman's constant,
T = absolute temperature,
μ = fluid viscosity,
d_p = particle diameter,
d_c = filter grain diameter,
V = filtration rate.

Particle transport can be influenced by media grain size, filtration rate, fluid temperature, and the density and size of the suspended particles (O'Melia and Stumm 1967), and this can be clearly seen within the three transport model equations. In fact, the dominant transport mechanism is dependent on particle size. For particles <1 μm in size, the more dominant transport mechanism is diffusion, whereas for particles >1 μm in size, the dominant mechanisms are sedimentation (Amirtharajah 1988) and interception (O'Melia 1985; Aim et al. 1997). As a result, there is a minimum net transport efficiency for particle sizes of approximately 1 μm (Habibian and O'Melia 1975; Amirtharajah 1988). In other words, particles that are about 1 μm in size are the least effectively removed particles by filtration, and this has been observed in a number of studies (Tobiason, Johnson and Westerhoff 1990; Tobiason et al. 1993; Aim et al. 1997).

Upon transport of the particle near the surface of the media grain, the particle must then attach to it. According to the DLVO theory, named after the work of Derjagin, Landau, Verwey, and Overbeek, a charged particle will have a layer of oppositely charged ions (the Stern layer) around it as well as a diffuse layer of mixed charged ions, which results in particles being repulsed from the media grain (Sawyer, McCarty and Parkin 1994). However, if the particle is destabilised enough from the coagulation unit process prior to filtration, electrostatic repulsions will be minimal allowing the particle to collide with the media grain surface. Through physical adsorption due to Van der Waals' forces, particles are thus able to attach to the grain or to other particles already attached to the media. Chemical adsorption may also play a role in particle attachment in which the particles become fixed to the surface of the media. Exchange adsorption is another possible form of interaction particularly for ions present in solution.

2.3.3 Filter Operation and Control

Typically at a WTP, the operator assesses the current state of the treatment process from data acquired from the Supervisory Control and Data Acquisition (SCADA) system or from information collected through testing. Then based on a series of jar tests and through operator experience, a control action is determined. The problem with this level of control is that with all the variables involved that influence filtration performance, it is difficult to completely assess the state of the system process and there is no definite clear-cut control action that can be taken. Moreover, jar tests only assess clarification performance, and therefore results from these tests do not necessarily reflect how well the filters will perform. The use of an ANN model to predict performance provides an alternate means of assessing the state of the treatment process as well as providing control options based on the determined relationships between variables and performance. Such determinations are more comprehensive than relying solely on operator assessment since the model is able to derive the extent of impact of multiple parameters making ANN assessments more reliable. Furthermore, current process control of treatment is a reactive process in which control actions are dictated and taken by changes in water quality coming off of the filters. However, since there is approximately a three to four hour detention time through the whole treatment process, the effects of control actions will not take effect for that time period, and this can potentially result in three to four hours worth of poor water coming from the plant before this is corrected. Utilising ANN models for process control is a proactive control technique in which water quality can be controlled based on influent water quality and operational control. Therefore, action is taken prior to incidences of poor water quality effluent by taking action in response to varying influent quality. This addresses the problem of the time delay of control actions resulting from reactive process control.

2.4 Factors Affecting Filtration Performance

Given an understanding of the filtration process as described in section 2.3, the factors that affect filtration performance can be determined and assessed, which is done in this section. This is necessary to determine the parameters that will be used in model development. These factors are categorised into three categories: operational parameters, solution chemistry and suspended particle properties.

2.4.1 Operational Parameters

2.4.1.1 Filtration Rate

According to models of transportation by sedimentation and by diffusion shown in section 2.3, it is indicated that particle collection is hindered by higher filtration rates (AWWA 1990). As can be seen in the equations, as filtration rate increases, efficiency of transport of particles to the filter media decreases, thus filtration performance decreases. However, there have been a number of studies done in the 1950's and 1960's on varying filtration rates and the subsequent effects on turbidity removal in which rates ranging from 5 m/h to 15 m/h were examined (AWWA 1990). It was found that the varying rates had little effect on effluent turbidity. Nonetheless, the conclusion made from this literature was that higher filtration rates result in slightly poorer filtrate, and it was shown that high rates up to 37 m/h had a detrimental effect (AWWA 1990). This result was also demonstrated in a different pilot plant study undertaken in Israel (Hatukai, Ben-Tzur and Rebhun 1997). It was shown that as filtration velocity changed from 20 m/h to 15 m/h, total particle counts decreased correspondingly, and it increased as filtration velocity increased back to 20 m/h. Interestingly, turbidity remained constant throughout the entire run. Such a result shows the sensitivity of particle counts compared to turbidity and implies that the effect of varying filtration rates may not be detected as well through turbidity. Therefore, the filtration rate may have a more significant effect on efficiency of particulate removal than what is indicated by earlier studies. This is further indicated in a bench scale study that was done in which the effects of filtration velocities of 0.11, 0.22, and 0.33 m/h on particle removal were examined (Moran et al. 1993). It was found that there was greater particle removal with a lower filtration velocity. Therefore, given the experimental evidence, filtration rate should be considered a significant factor affecting filtration performance.

2.4.1.2 Filter Media Size

It is generally understood that the larger the media, the less efficient the particle removal. This is because with increasing diameter of the filter particles, less surface area is available for contact to be made with particles, thus less opportunity for particles to be retained. This is supported theoretically by the model of efficiency of transport by interception (eqn. 2) and by the model for transport by diffusion (eqn. 3) given in section 2.3.2. From these equations, it can be seen that as filter grain diameter decreases, efficiency of transport of particles increases resulting in improved filtration performance. Experimental evidence from a bench scale study supports this statement (Moran et al. 1993). 1.85 mm media was tested against 0.78 mm media with the same depth of 746 mm. It was found that clean bed removal was more efficient for the smaller 0.78 mm media than the 1.85 mm media. Moreover, when adjustments were made to filter depth so that surface area was constant for both media sizes, particle removal efficiencies were very similar. This supports the concept that the amount of surface area of filter media influences efficiency of particle removals. Therefore, filter media size is a significant factor in filtration performance.

2.4.1.3 Filter Media Depth

Little information was encountered in literature regarding effects of filter media depth on filtration. However, it is intuitive to deduce that increasing the depth of the filter would improve filtration. This is because increasing filter depth increases total filter media surface area thereby allowing for greater opportunity for particles to have contact with the filter media. As a result, particle removal efficiency should improve. A bench scale lab study was done in which particle removal was determined at filter depths of 193 mm and 746 mm (Moran et al. 1993). It was observed and concluded that particle removal increased with increasing depth. Therefore, filter depth can be considered to be an important factor in filtration performance.

2.4.2 Solution Chemistry

2.4.2.1 Coagulant Dosage and Type

The two general mechanisms that occur in filtration are the transport of particles to the filter grain followed by adherence of the particle to the grain. In order for particles to attach, they must be chemically destabilised by pre-treatment, and the importance of pre-treatment has been emphasised in literature (O'Melia 1985; Amirtharajah 1988). Given that coagulants are responsible for destabilisation, it is clear that coagulant dosage is a significant factor in filtration efficiency.

A series of pilot plant experiments were performed which demonstrates the importance of coagulant dosage to filtration performance (Amirtharajah 1988). Alum dosages of 1, 5, and 8 mg/L were examined, and an optimal alum dosage was observed at 5 mg/L at a pH of 6.9-7.0. Good filtration was also observed at a dosage of 8 mg/L, however poor turbidity removal occurred with a dosage of 1 mg/L.

In a different lab study (Tobiason and O'Melia 1988), calcium nitrate was used as a coagulant in varying concentrations to observe the clean-bed particle removal of a monodisperse solution of 4 μm latex particles. It was found that as concentration of the coagulant increased, particle removal increased further demonstrating the influence of coagulant dosage. Zeta potential, which is explained in section 2.4.3.1 on p. 34, was also measured in this experiment, and it was observed that as coagulant dosage increased, the zeta potential of the particles approached 0 mV compared to an initial negative value of ~70 mV. This is an indication of how the coagulant destabilises the particles eliminating the negative charge.

Coagulant type can also be considered to be an influential factor on filtration performance as some coagulants may be more effective than others at destabilising particles depending on solution conditions. An example of this was demonstrated in a study in which alum and $Fe_2(SO_4)_3$ were compared (Valade et al. 1996). It was observed that the iron sulfate produced slightly lower filter effluent turbidity than alum, which is evidence that coagulant type is an important factor that can influence filtration performance.

From these studies, it is apparent that coagulant type, and more particularly the dosage, is very important in particle removal. Coagulant dosage should definitely be considered as a significant factor in filtration performance.

2.4.2.2 Polymer Dosage and Type

Polymers can be used as a sole coagulant, coagulant aid or filter aid. It differs from metal salt coagulants, such as alum, in that polymers are long chain organic molecules with positive and/or negative charges and may destabilise particles possibly by charge neutralisation or by interparticle bridging. Therefore, the effects of polymer characteristics may not be as significant depending on the dominating mechanism. This was demonstrated by a pilot plant study that was done in which the effects of using polymers as a filter aid on filtration performance were assessed (Zhu et al. 1996). Different molecular weight polymers with different charge densities were examined, and it was found that these parameters had little effect on filtrate quality. However, since the polymers were used as filter aids and used in small amounts, it is difficult for charge neutralisation to occur, therefore polymer type and charge density would have little impact on filtrate quality (Zhu et al. 1996). It is more likely that the dominant mechanism of removal was interparticle bridging between the polymers and particles, which would explain the observed lack of influence of polymer type. Polymer dose, however, was found to be an important factor. An optimal dose was found to be 0.01 mg/L with no improvement when dose was increased to 0.02 mg/L. However, when dose was decreased to 0.005 mg/L, filtrate quality degraded.

In another pilot plant study (Hatukai, Ben-Tzur and Rebhun 1997), it was shown that addition of cationic polymer to alum as a coagulant aid resulted in better turbidity and total particle count removal. Log removal increased from 1.4 to 1.8 for particles in the size range of 2-5 μm with the addition of 0.1 mg/L of polymer. Furthermore, increasing the polymer dosage from 0.1 mg/L to 0.3 mg/L also improved removal efficiency. There was less improvement in the removal of smaller sized particles, but there was good improvement in the removal of coarser particles at this higher dosage. This is another indication on the importance of polymer dose to particle removal efficiency.

It was shown in the study by Zhu et al. that polymer type was not found to be a significant factor in filtration performance. However, both studies discussed have shown polymer dose to be an influential factor and therefore should be considered as significant in filtration performance.

2.4.2.3 pH and Alkalinity

One of the main effects of pH during water treatment is on the solubility and speciation of the coagulant alum. This in turn affects the degree of destabilisation of particles as well as determines the dominant mechanism(s) of destabilisation, which ultimately affects filtration performance. The relationship between pH and alum concentration is shown in Figure 2-1, which clearly displays solubility of alum at a given pH. The zeta potential of a colloid in a solution of 10 mg/L of alum and without alum is compared in the lower half of the diagram. pH can also influence the surface charge of a particle (Stumm and Morgan 1996). At low pH, protons can adsorb to the surface of a negatively charged particle. Conversely, at high pH, hydroxides can adsorb to the surface of a positively charged particle. Either way, the surface charge is affected and the particle can become destabilised.

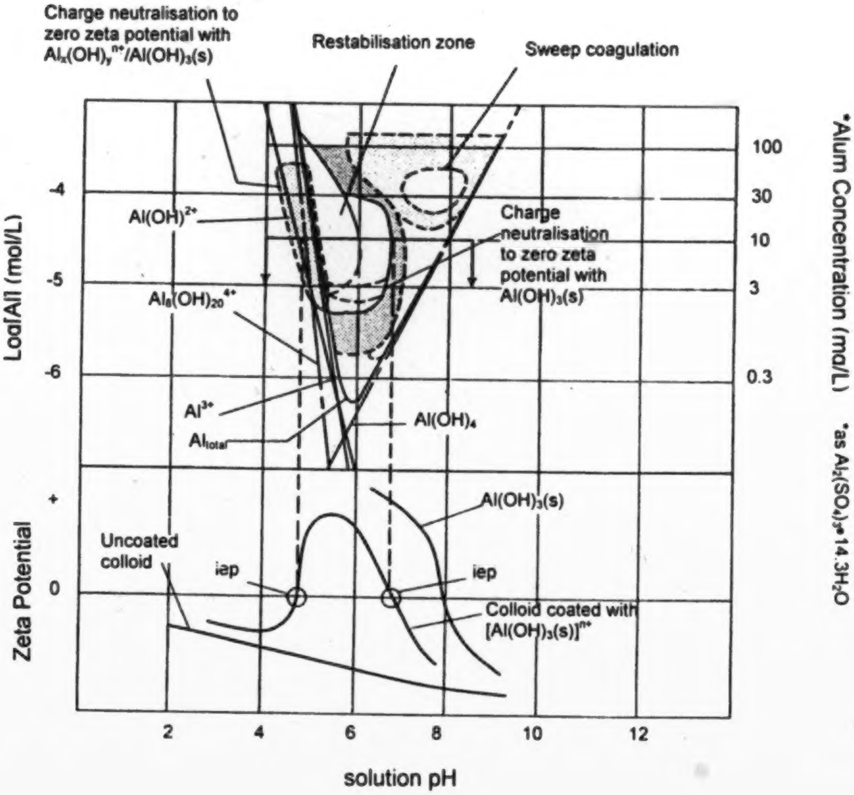


Figure 2-1. Alum coagulation diagram and zeta potential with charge neutralisation zones (modified from Amirtharajah (1988)).

In one direct filtration bench study (Collins, Amy and Bryant 1987), the effect of varying pH on filtration was determined. It was observed that an initial pH of 8.5 generally resulted in lower final turbidity and total particle counts compared to an initial pH of 5.5, but a clear explanation of these results was not discussed. The problem lies in the fact that alum dosages were chosen based on near-maximum removal of non-volatile total organic carbon. This resulted in holding alum of dosage to 2.0 mg/L for pH 5.5 and 4.0 mg/L for pH 8.5. Therefore, the reason for the lower turbidity and particle count at pH 8.5 compared to 5.5 was likely due to the higher amount of alum present whose actions result in improved filtration.

The opposite results were found, however, in another lab study (Tobiason and O'Melia 1988). In this study, the zeta potential of a monodisperse mixture of 12 μm latex particles was measured as pH was changed. It was observed that as pH decreased from 10 to 2, zeta potential changed from –55 mV to ~ 3 mV. Filter efficiency was then observed for pH 9.7 and 3.0, and it was found that particle removal efficiency was better at pH 3 than at 9.7. This observation is explainable by the likely fact that at such a low pH, protons adsorbed to the surface of the particle, which accounts for the observed zeta potential change to a positive value at lower pH. As a result, the particles are destabilised allowing for greater interaction and attachment with the negatively charged filter media. At a pH of 9.7, the particles would tend to retain a highly negative charge and therefore would be less likely to interact and be retained by the filter media. Therefore, the filter efficiency would be greater at a pH of 3.0 than at 9.7.

Another study evaluated the effect of pH on direct sand filtration at pH values of 5, 7, and 9 (Ebie and Miyake 1991). It was found that turbidity removal was higher at pH 5 and 7 compared to pH 9, however the mode of filtration was different at each pH. At pH 5, particles had a highly positive charge and were therefore retained in the filter media mainly due to electrostatic attractive forces. At pH 7, which is in the optimal coagulation range of kaolinite particles with alum, larger flocculated particles formed with the predominant transport mechanism found to be interception. This is different from filtration at pH 5 because the effectiveness of filtration is due to both particle attachment and transport, whereas at pH 5, filtration efficiency was mainly due to particle attachment. At pH 9, the suspended particles and the filter sand grains had a negative charge, and it was assumed that particles passed directly through the filter.

Another study found that optimal filtration occurred at a pH range of 6.9-7.1 with a sufficient dosage of alum for destabilisation (Amirtharajah 1988). This range of pH also falls within the pH boundary for charge neutralisation.

It has been shown through these studies that pH has a major impact on filtration performance. It influences alum solubility therefore affecting the degree of particle destabilisation and subsequently particle attachment to filter media. pH also impacts filtration by affecting the surface charge of the particle with the adsorption of protons onto the negatively charged surface of the particle resulting in particle destabilisation. pH should therefore be considered as a significant factor in filtration performance.

Alkalinity is defined as the measure of the capacity to neutralise acids (Sawyer, McCarty and Parkin 1994). This is an important characteristic of water because it has an effect on pH in that it would influence how much pH decreases as acid is added to water. Such a situation arises when alum is added to water. As alum undergoes hydrolysis, a proton is released which results in a decrease in pH. The alkalinity of the water, however, would determine the extent to which pH is affected. Since changes in pH have an effect on filtration performance and alkalinity affects changes in pH, alkalinity can be said to have an effect on filtration performance and may be a significant factor.

2.4.2.4 *Ionic Strength*

The ionic strength of the solution may influence the surface chemistry of the particles and alter the interactions between particles and between particles and filter media, which consequently may have an effect on filtration performance. If ionic strength is sufficiently high enough, the double layer of the particle is compressed with the diffuse layer reduced (AWWA 1990). This would reduce the range of repulsive interaction between similarly charged particles allowing the particles to interact more easily through van der Waals interactions. A lab study was done (Prasanthi et al. 1994) with hematite particles in which the ionic strength was varied from 10^{-3} to $8x10^{-2}$ M by varying the KCl concentration from 0 to 80 mM. With increasing ionic strength, colloidal aggregate size increased from 0.085 um to 0.99 μm. Clean bed filter efficiency thus increased with increasing size within this size range. However, efficiency decreased as aggregate size neared 1 μm, which is theoretically expected as discussed in section 2.3. During the transient stage of filtration, there was quick breakthrough with KCl concentrations from 0 to 10 mM. For higher concentrations (40 to 80 mM), initial removal efficiency improved with decreasing concentrations, however the trend was reversed as breakthrough was approached in that removal efficiency improved with increasing concentrations. Based on these experimental results, ionic strength can have an impact, and it may be considered a significant factor.

2.4.2.5 Particle Concentration

Influent particle concentration is an important parameter in that it has an influence in deciding the type of filtration process to utilise. Depending on the mass concentration and volume average diameter of particles in the raw water, there are three types of filtration processes to consider as an optimal treatment process: contact filtration, direct filtration, and conventional filtration (O'Melia 1985). This is shown in the Figure 2-2 below. Generally, for larger sized particles in greater concentrations, conventional filtration is most effective.

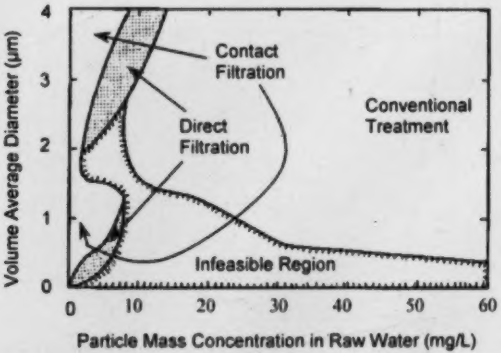


Figure 2-2. Optimal water treatment configuration as a function of raw water characteristics (modified from O'Melia (1985)).

More importantly is the effect of particle concentration on filter performance. Given that the purpose of filtration is to remove particles, variations in influent particle concentration have the potential to affect performance. Such effects are related to particle size in that the presence of different sized particles also affects performance. Thus, varying concentrations of different sizes of particles have a varying effect on filter efficiency (Prasanthi, Vigneswaran and Dharmappa 1997). This is partially due to the different mechanisms that dominate depending on the particle size as discussed in section 2.3.

The removal efficiency of a clean filter is independent of influent concentration, but as filtration proceeds, concentration is observed to have a significant effect (O'Melia and Ali 1978). Specifically, it was observed that as influent concentration of suspended particles increased from 1 mg/L to 50 mg/L, the percent concentration remaining decreased. It was also found in the same study that during the filter ripening stage, filter efficiency was higher for higher influent concentrations. In another study in which submicron particles of sizes 0.46 µm and 0.825 µm were used, it was observed that the clean bed filter efficiency increased with increasing influent mass concentration (Prasanthi, Vigneswaran and Dharmappa 1997). Given the observations made in these studies, particle concentration should be considered to be a significant parameter in filtration.

2.4.2.6 Hardness and Water Softening

Hardness of water is caused by multivalent metallic cations and is derived largely from contact with soil and rock formations (Sawyer, McCarty and Parkin 1994). Some of the principal cations that cause hardness in water include Ca^{2+}, Mg^{2+}, Sr^{2+}, Fe^{2+}, and Mn^{2+}. Very little information was encountered in the literature regarding the effects of hardness on filtration performance. It was stated, however, that the colloidal stability of natural particles in surface water depends significantly on hardness and dissolved organic carbon or humic substances in the water (Amirtharajah 1988). This in turn can affect the ability of the particles to interact with the filter media and each other implying that filtration performance could ultimately be affected.

In some treatment plants where the source water is quite hard, hardness is removed by softening. Lime is a typical chemical added to soften the water, and this results in the formation of calcium carbonate, which precipitates and settles out. This reaction also results in an increase in pH, which is then controlled. Because of this process, the formation of calcium carbonate precipitate may add to the particle concentration that enters the filters, which in turn could affect filtration performance as discussed previously in section 2.4.2.5. Therefore, it is possible that if water softening by lime addition is performed at a treatment plant, lime dosage could be considered to be an influential factor in filtration performance.

2.4.2.7 Organic Matter Concentration

Organic matter is capable of adsorbing onto particles, which can alter the surface characteristics of the particle thus altering interactions with other particles (Prasanthi et al. 1994). In fact, colloidal stability of natural particles in surface water is dependent on dissolved organic carbon or humic substances in water (Amirtharajah 1988). Since the surface chemistry of particles and the extent of particle stability affect particle removal, filtration performance may be affected by organic matter concentration.

A lab study was done (Prasanthi et al. 1994) with hematite particles in which fulvic acid concentration was varied from 0 to 4 mg/L to determine the effect on particle removal. As concentration increased from 0 to 0.5 mg/L, the zeta potential of the particles went from a positive charge to a net charge of near zero, and the particle aggregate sizes increased. Clean bed filter efficiency was highest at a concentration of 0.5 mg/L fulvic acid. Removal efficiency dropped at 0.75 mg/L fulvic acid because aggregate size was at approximately 1 um, which is the critical size in which particle transport is at a minimum. At 2 and 4 mg/L fulvic acid, there was virtually no aggregation because the particles were stabilised by the fulvic acid as evidenced by the particles possessing a negative zeta potential. However, removal efficiency was still quite high due to increase in transport of particles from diffusion.

In another study that was done Collins et al. (1987) found that the presence of higher levels of fulvic acid did not significantly affect final turbidity and total particle counts at a pH of 8.5. However, at pH 5.5, inhibition of particle removal was observed.

The results from these studies are evidence that organic matter can interact with and affect the chemistry of particles. Consequently, organic matter concentration might be considered to be an influential factor in filtration performance.

2.4.2.8 Activated Carbon Dosage

Powdered activated carbon (PAC) is predominantly used for controlling taste and odour by adsorbing organic compounds. By affecting the concentration of organic matter in water, filtration performance may subsequently be affected as discussed previously in section 2.4.2.7. Furthermore, addition of activated carbon to water may contribute to particle concentration that eventually enters the filters. As previously mentioned in section 2.4.2.5, the added particle concentration may impact filtration performance. Given the effects that activated carbon potentially has, it should be considered to be an influential parameter in filtration performance.

2.4.2.9 Temperature

From the models shown for transport efficiency by sedimentation (eqn. 1) and diffusion (eqn. 3) in section 2.3.2, it can be seen that transport is dependent on temperature for diffusion and viscosity for both sedimentation and diffusion. As temperature decreases, efficiency of transport by diffusion would decrease. In addition, viscosity increases with decreasing temperature, which results in a decrease in efficiency of transport of particles by sedimentation and diffusion. This can ultimately result in a decrease in particle removal efficiency in filtration. Settling in the clarification tanks can also be affected. Another effect that temperature can have is on reaction rates. Cold temperatures can result in a decreased reaction rate affecting upstream processes such as coagulation and softening. Consequently, if the extent of destabilisation in coagulation is decreased, filtration efficiency may decrease. Therefore, temperature should be regarded as an important factor in filtration performance. However, the impact may only be slight if temperature does not vary greatly.

2.4.3 Suspended Particle Properties

2.4.3.1 Zeta Potential

Zeta potential can be described in the following way. A charged particle (negative) accumulates a layer of ions around it, which is the diffuse layer, with ions of the opposite charge (positive) particularly around the surface to form a layer called the stern layer. Together, these layers comprise the electric double layer, and this is shown in Figure 2.3. The potential at the edge of the stern layer within this double layer is the zeta potential (Sawyer, McCarty and Parkin 1994). The zeta potential is dependent on the potential at the surface of the particle, which is the Nerst potential, and the thickness of the double layer, and its value determines the extent of the electrostatic forces of repulsion between charged particles (AWWA 1990). Particles in natural water typically have a zeta potential ranging from –20 to –40 mV, but suspensions of particles that are well destabilised have a potential close to 0 (AWWA 1990). Therefore, it is evident that zeta potential is a good indication of the extent of particle destabilisation, which is an important condition in filtration effectiveness.

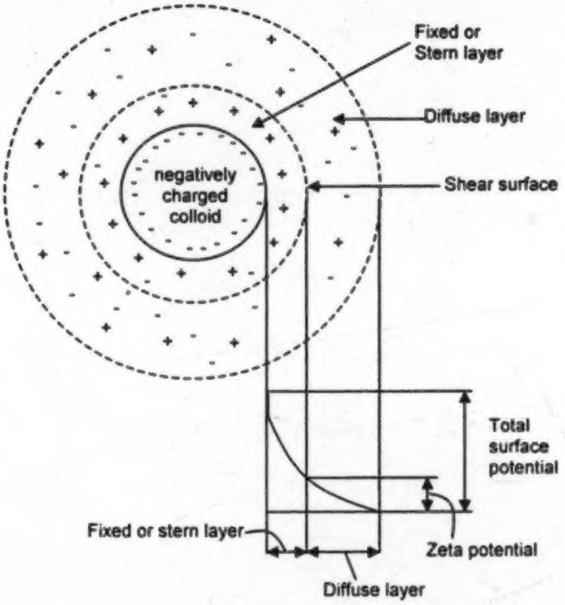

Fixed or
Stern layer
Diffuse layer
Shear surface
negatively
charged
colloid
Total
surface
potential
Fixed or stern layer
Zeta potential
Diffuse layer

Figure 2-3. Electrical double layer of a negatively charged colloid (adapted from Sawyer, McCarty and Parkin (1994)).

In a study previously mentioned in the section on coagulant dosage in section 2.4.2.1, zeta potential was monitored as coagulant dosage varied (Tobiason and O'Melia 1988). It was observed that as coagulant dosage increased, the zeta potential of the particles approached zero, whereas the particles initially possessed a negative zeta potential. As well, as zeta potential approached 0 mV, removal efficiency improved significantly. Within the same study, another experiment was carried out in which zeta potential was monitored while the pH was varied. What was observed was that as pH was shifted from 10 to 2, zeta potential of the particles changed from –55 mV to nearly 0 mV. Furthermore, as zeta potential became less negative, removal efficiency improved. These results indicate a correlation between zeta potential and filtration performance, thus it may be an important factor.

2.4.3.2 Particle Size and Size Distribution

Particle size is an influential factor in filtration performance in that the dominant mechanism of transport of particles to filter media is dependent on the size of the particle. As it was discussed in section 2.1, there is a minimum net transport efficiency for particle sizes of approximately 1 μm (Habibian and O'Melia 1975; Amirtharajah 1988). Particles that are about 1 μm in size are least effectively removed by filtration, and this was observed in the following studies.

In one lab study of contact filtration in which monodisperse suspensions of 0.27, 1.32 and 10.0 μm particles were used, it was found that the 1.32 μm particles were removed least efficiently. Removal of the submicron particles was better, and removal of the large 10 μm particles was best (Tobiason, Johnson and Westerhoff 1990; Tobiason et al. 1993).

In another granular bed filtration study (Aim et al. 1997), monodisperse suspensions of 0.46, 0.825, and 2.967 μm latex particles were studied. It was observed that the clean bed filter efficiency was less for the 0.825 μm particles than the other two sizes, which supports the contention that minimal particle removal occurs for particles approximately 1 μm in size. However, it was also found that such a critical size did not exist for the ripening stage filter efficiency and that removal efficiency at this stage increased with particle size.

Size distribution may also play a role in filtration efficiency in that particles of one size can affect the removal efficiency of particles of a different size (Aim et al. 1997). In the contact filtration study (Tobiason et al. 1993), the removal efficiency of a polydisperse suspension of 0.27 and 1.32 μm particles was examined. It was found that for the 0.27 μm particles, the presence of the larger particles resulted in a decrease in clean bed removal efficiency of the smaller particles. On the other hand, for the 1.32 μm particles, the presence of the smaller particles resulted in an improvement in clean bed filter efficiency of the larger particles. The improvement in filter efficiency for larger particles due to smaller particles could have been because of increased apparent surface roughness of the media or of the larger particle due to small particle deposition. Conversely, the effect of larger particles on smaller particles could be due to unfavourable hydrodynamic interactions or differences in destabilisation (Tobiason et al. 1993). These effects of one particle size on another are dependent on the relative abundance of each particle as well as chemical conditions (Tobiason, Johnson and Westerhoff 1990).

Given that transport mechanisms are size dependent and that there is experimental evidence supporting this, it is apparent that particle size is a significant factor that affects filtration performance. As well, particle size distribution can also be considered a significant factor in filtration performance from observations made in lab studies.

2.5 Aspects of Artificial Neural Networks

The purpose of this section is to discuss the basic fundamentals of ANN. It is important to have an understanding of ANNs in order to be able to design, develop and optimise an ANN model. A description of ANN is presented in this section along with its specific components and aspects.

2.5.1 Description

An artificial neural network is a biologically inspired computational model (Kasabov 1996; Garrett Jr., Gunaratnam and Ivezic 1997). It is classified as an artificial intelligence modelling (AI) technique in that it has the capability to learn from examples, which is a key characteristic. An example of a basic network is depicted in Figure 2-4. The network consists of highly interconnected processing units called neurons, which are the basic elements of an ANN. These processing units or neurons receive input signals

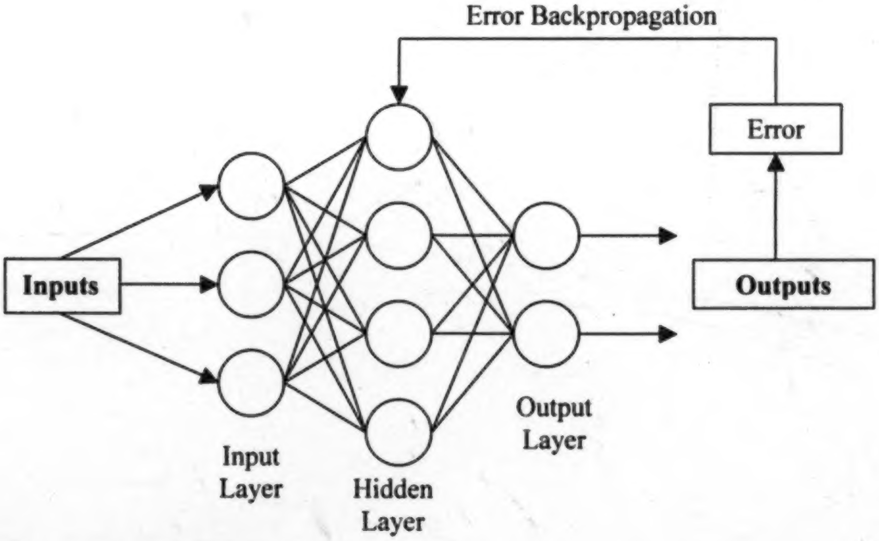


Figure 2-4. A basic artificial neural network.

from other neurons or external stimuli, process it through an activation or transfer function, and produces a transformed output signal to other neurons or external outputs (Zhang, Patuwo and Hu 1998). Learning occurs through training of the network in which input examples are repeatedly presented to the network, and modification and optimisation of the connection weights between neurons are performed through the use of a learning algorithm. Therefore, mathematical formulae or algorithms are not required to develop solutions to the problem. Other important characteristics are its robustness and its ability to generalise rules and apply them to new cases. The network is able to produce the best output according to training examples when new input vectors are presented to the network, and it is fault-tolerant in that the system is still able to perform well if there are errors within the network (Kasabov 1996). In other words, approximately correct answers are produced even though data presented to the network is partially incorrect or incomplete.

Before ANN can be applied to a problem, it should be determined if using ANN is appropriate and applicable. ANN is applicable in cases in which the algorithms or heuristics to solve the problem are unknown or too expensive or difficult to determine. In addition, developing mathematical formulae for the case study should not be a goal since ANN does not use nor provide such formulae. Furthermore, success of the model is dependent on data as the modelling process is data intensive. If an adequate amount of data are not available or if the type of data available is insufficient to describe the problem domain, then ANN may not be applicable. It is also important to have expert knowledge of the study domain itself. In the case of filtration performance, it is known that this is a complex process with multiple factors influencing the process. The algorithms and heuristics to define and describe all the complex relationships involved are not completely known and are quite difficult to determine. Since mathematical formulae are not necessary for this study, and provided that there is sufficient variety and quantity of data available that will represent the domain of this process, using ANN is applicable to the modelling of filtration performance in a WTP.

ANN has previously been applied to other environmental engineering problems. Such applications have included water demand forecasting, river water quality prediction, air quality prediction, pipe break analysis, and enhanced coagulation modelling.

2.5.2 Architecture, Components, and Aspects of ANN

There are several main aspects of ANN models: a set of processing units called neurons, a state of activation for each neuron, a set of connection weights, a propagation rule, an activation rule, and a learning rule (Rumelhart and McClelland 1986).

2.5.2.1 Neurons

Neurons are processing units that receive input signals, process this information and then transmit an output signal. There are three types of neurons in a neural network: input, output, and hidden neurons. Input neurons receive input from external sources, which is scaled to a value to between 0 and 1 or –1 and 1 depending on the type of scaling function used. This scaling function is present only within input neurons. Their output is then computed as a function of the activation level, which is then transmitted to the network (Garrett Jr., Gunaratnam and Ivezic 1997). Specific examples of inputs from models developed for this thesis include raw water parameters such as turbidity and temperature and operational parameters such as alum dose or polymer dose. Output neurons receive input from the rest of the network. The output of these neurons are computed and sent out of the system (Garrett Jr., Gunaratnam and Ivezic 1997). An example is effluent particle counts, which was used in model development. Hidden neurons are those units in which inputs are received and outputs are sent within the network with no external contact (Garrett Jr., Gunaratnam and Ivezic 1997). These neurons are very important in that they allow the network to detect and capture features and patterns in data and to perform the non-linear mapping between input and output variables (Zhang, Patuwo and Hu 1998).

As it was introduced in section 2.5.1, neurons process information in the following way. Input signals from other neurons or external stimuli are received by a neuron. This would include signals from error backpropagation as well. This is then processed through an activation or transfer function producing a transformed output signal that goes out to

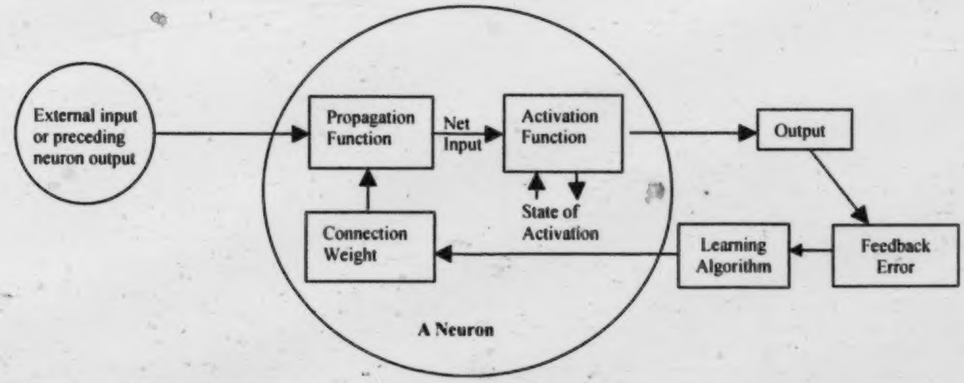

External input
or preceding
neuron output
Propagation
Function
Net
Input
Activation
Function
Output
State of
Activation
Connection
Weight
Learning
Algorithm
Feedback
Error
A Neuron

Figure 2-5. A basic neuron.

other neurons or as external outputs. A diagram of a neuron is shown in Figure 2-5.

2.5.2.2 State of Activation

The state of activation of a neuron determines whether or not the neuron will fire off an output. This is determined by the neuron after processing input stimuli and calculating whether or not a certain threshold is exceeded by the transformed signal.

2.5.2.3 Set of Connection Weights

Neurons are connected to each other and communicate through these connections. The strengths of these connections determine how the network will respond to a given input (Rumelhart and McClelland 1986; Garrett Jr., Gunaratnam and Ivezic 1997). There are three types of connections: excitatory, inactive, and inhibitory (Garrett Jr., Gunaratnam and Ivezic 1997). For example, if the weight, w_{ij}, of the connection from the ith neuron to the jth neuron in a network with a sigmoidal activation function is positive, the connection is excitatory meaning that neuron i has an activation effect on neuron j. If w_{ij} is zero, then the connection is inactive meaning that neuron i has no effect on neuron j. If w_{ij} is negative, then the connection is inhibitory meaning that neuron i hinders activation of neuron j.

2.5.2.4 Propagation Function

The rule of propagation determines the net input to a neuron by combining the output of a preceding neuron and the connection weight. The total net input from all preceding neurons can be mathematically represented by (Garrett Jr., Gunaratnam and Ivezic 1997):

$$N_j = \sum_i w_{ij} o_i \qquad (4)$$

where,

N_j = total net input,

w_{ij} = connection weight,

o_i = output of preceding neuron

2.5.2.5 Activation Function

The activation function, which is also referred to as the transfer function, determines the new level of activation of a neuron by combining the total net input received with the current level of activation. This can be represented mathematically by (Garrett Jr., Gunaratnam and Ivezic 1997):

$$a_{inew} = F\left[a_{iold}, N_i\right] \qquad (5)$$

where .

a_{inew} = new activation level,

F = activation function,

a_{iold} = current activation level

N_i = net input

The current activation level acts as the threshold value to which the new activation level is compared as mentioned in section 2.5.2.2.

2.5.2.6 Learning Rules and Network Learning

In order for the neural network model to learn the relationships between inputs and outputs, the pattern of connectivity in a model would need to be modified. The learning rule determines how the network is modified (Garrett Jr., Gunaratnam and Ivezic 1997). There are three types of modifications that can be done: the development of new connections, the loss of existing connections, or the modification of the strengths of connections that currently exist (Rumelhart and McClelland 1986). The development and loss of connections can be considered to be special cases in that connection weights become zero for a connection to be lost or move away from zero for a connection to be gained (Rumelhart and McClelland 1986). Most neural network simulators, however, only provide the capability to modify existing connection strengths (Garrett Jr., Gunaratnam and Ivezic 1997). In which case, a zero connection weight indicates an inactive connection.

Most learning rules involving the modification of current connection weights are variants of the Hebbian learning rule. The general idea of this rule is that if a neuron, u_i, receives input from another neuron, u_j, and both neurons are highly active, then the connection weight, w_{ij}, from u_i to u_j should be strengthened (Rumelhart and McClelland 1986).

Another learning rule, which is utilised in the models developed in this study, is the delta rule. Basically, an input is used by the model to generate an output and this is compared with the desired or actual output. Connection weights are then modified to minimise the squares of the difference between the generated output and the actual desired output.

There are two phases in the application of the generalised delta rule (Rumelhart and McClelland 1986). First, input is presented to the network and propagated forward through the network to generate an output value. This output is then compared with the actual output, and an error signal δ is computed from the difference between the outputs times the derivative of the squashing function. In the second phase, δ is propagated backwards through the network, and appropriate weight changes are made. Mathematical derivation of the generalised delta rule is presented by (Rumelhart and McClelland 1986).

Supervised learning is a type of learning that utilises the delta rule. Training examples include both the input vectors and the desired output vectors. Training is performed until the neural network is able to associate each input vector with the corresponding desired output vector (Kasabov 1996). Unsupervised learning is a different type of learning. It differs from supervised learning in that the training examples only include the input vectors and not the output vectors. The neural network learns the internal features of the input vectors presented to it and categorises outputs according to how they associate with each other (Kasabov 1996). Kohonen categorisation is associated with this type of learning and is further discussed in section 2.5.4.

2.5.3 Generalisation and Convergence

Convergence is the ability of the network to learn the training data within the specified error tolerance, whereas generalisation is the ability of the network to produce reasonable results for the unknown situation after training has been completed. This is important particularly during model development because it describes the ability of the network to learn and then perform. The two main reasons for a network to not generalise or converge well are because the data set is too noisy or is not representative enough of the problem domain. Noise is further discussed in section 3.2. Another reason for poor generalisation includes previously unknown data deviating too much from examples used in the learning (Muller and Reinhardt 1991). By avoiding these conditions, model development should be easier to accomplish.

2.5.4 Kohonen Categorisation

As it will be discussed in the methodology section in section 3.0, Kohonen networks were used as part of the modelling process in this study. Therefore, the concept of this network is described here to provide an understanding of what Kohonen networks are and how they work in order to understand how and why they are utilised.

Kohonen networks are based on the behaviour of biological neurons in which specific neural responses are associated with the spatial location of neurons. The network comprises an interconnected lattice or array of neurons each of which are connected to input neurons. As network training occurs, example inputs are presented to the network, and neuron connection weights are adjusted. The output neuron and its connection weights, which is equivalent to the centre of the category, that most resembles the input is the "winner" and its weights are changed accordingly. In addition, neurons in the neighbouring vicinity to the winning neuron are also adjusted. As a result, those input patterns that have similar characteristics tend to activate neurons that are clustered together. The resulting output is the categorisation of each input data pattern. Thus, significant features and patterns within the training data influence the network and are categorised accordingly, and this type of network learning is defined as unsupervised learning. This differs greatly from backpropagation networks, which employ supervised learning and requires the actual outputs in addition to the inputs. In other words, the correct output is not required in the training of Kohonen networks, only inputs are needed, whereas the actual outputs are required to train backpropagation networks.

Data patterns are categorised based on the spatial relation of the pattern to the centre of a category in n-dimensional space in which n-dimensional space is defined by the number of inputs. The centre of the category is defined by the average value of the inputs of all patterns within that category. The spatial distance of the pattern is calculated based on the input values and relative to the centre of the category in terms of the Euclidean distance, which is defined by equation 6:

$$\|X - Y\|_2 = \left\{\sum_{(n=1,N)} (X_n - Y_n)^2\right\}^{1/2} \tag{6}$$

where, X is the input value,

Y is the value of the centre, and

n is the number of inputs that define the pattern.

The smaller the Euclidean distance, the closer a pattern is to the centre of a category, and the more likely it is that the pattern belongs to that category.

2.6 Other AI Systems

ANN is not the only type of AI modelling technique that is available. These other AI techniques have been applied to water treatment (Boscolo, Mangiavacchi and Drius 1991; Zhu and Simpson 1996; Liu and Wu 1997; Evans et al. 1998). However, there are advantages and disadvantages of each technique that need to be considered. Ultimately, upon evaluation and given the available resources as well as considering previously successful application of ANN in the environmental engineering field, which includes water demand prediction (Stark, Stanley and Buchanan 2000), prediction of cast iron water main breaks (Sacluti, Stanley and Zhang 2000), air quality prediction (Hasham, Stanley and Kindzierski 1998), modelling of coagulant dosages (Mirsepassi, Cathers and Dharmappa 1995; Gagnon, Grandjean and Thibault 1997; Han et al. 1997), pressure drop in filtration (Conlin, Peel and Montague 1997), chlorination dosage and control (Rodriguez and Serodes 1996; Rodriquez, Serodes and Cote 1997), enhanced coagulation application (Stanley and Zhang 1997), and water quality parameter prediction (Maier and Dandy 1996; Zhang and Stanley 1997), ANN is still the modelling technique of choice. There has been little research, however, in the modelling of the treatment process of an entire WTP. The examples of applications in water treatment given here are the modelling of single unit processes in treatment. Therefore, development of an ANN model encompassing the entire treatment process, as with the case in this study, is a unique approach toward modelling water quality from a plant.

2.6.1 Expert Systems

An expert system is a computer system that simulates human experts in a given area of specialisation (Castillo, Gutierrez and Hadi 1997). Human reasoning is simulated about a problem domain with reasoning performed over representations of human knowledge, and problems are solved by heuristics or approximate methods (Jackson 1990).

The basic architecture of an expert system is made up of six components: knowledge base, context, inference mechanism, explanation facility, knowledge acquisition, and user interface (Maher and Allen 1987). The knowledge base contains the facts and heuristics or rules associated with the problem domain, and the context contains information about the problem being solved. The inference mechanism is the heart of an expert system, which contains the control information and uses the knowledge base to modify and expand the context. The explanation facility is the component that can trace the execution of the system program and provide the explanations behind a solution to the problem being solved. The knowledge acquisition facility is the component that facilitates entering knowledge into the knowledge base. Finally, the user interface is the liaison between the expert system and the user where the user can interact with the system. Additional subsystems include coherence control, information acquisition, action execution, learning (Castillo, Gutierrez and Hadi 1997). The coherence control subsystem controls the consistency of the knowledge base and prevents incoherent knowledge from reaching the knowledge base. Information acquisition is utilised by the inference engine when initial knowledge is limited and conclusions cannot be made. The user can provide such information through the user interface. The action execution subsystem allows the expert system to perform actions based on conclusions or solutions made. Expert systems can learn and gain experience based on available data by storing knowledge, which can be drawn upon by the knowledge acquisition subsystem.

Compared to algorithmic solutions, heuristics are not guaranteed to succeed, and the system must be capable of explaining and justifying solutions or recommendations (Jackson 1990). However, mathematical algorithms might be too difficult to develop in some cases, whereas rules for expert systems are easier to develop. In addition, expert systems can acquire knowledge as well as verify its coherence, store the knowledge or ask for new knowledge when needed, learn from the knowledge base and available data, reason and make inferences in deterministic and uncertain situations, and communicate with experts, non-experts, and other expert systems (Castillo, Gutierrez and Hadi 1997). The main problem with utilising an expert system, however, is that the general rules developed might be too simplistic to describe more complex problem domains. For example, filtration performance at a WTP involves many complex processes, therefore the rules that would be developed for an expert system may not adequately describe all the relationships involved that affect filtration performance. As well, experts would need to be available to develop and modify rules as needed.

2.6.2 Fuzzy Logic

Fuzzy systems are rule-based expert systems based on fuzzy rules and fuzzy inference, which represent knowledge that is subjective, ambiguous, vague, or contradictory (Kasabov 1996). There are three main components in a fuzzy system: fuzzy input and output variables defined by their fuzzy values, a set of fuzzy rules, and a fuzzy inference mechanism. Fuzzy propositions are used in that propositions contain fuzzy variables and values such as "the temperature is cold" or "the pH is high." Fuzzy inference takes inputs, applies fuzzy rules, and produces outputs, and inputs and outputs may be either crisp exact values such as "2" or "3," or fuzzy values such as "low" or "high" (Kasabov 1996). Defuzzification is then done in which an output membership function is transformed into a single value.

Fuzzy logic has an advantage over expert systems in that rule generation can be done with statistical tools using historical data rather than relying on experts themselves. Another advantage just as with expert systems is that mathematical equations are not necessary to solve the problem. Other advantages of fuzzy systems are that they are easy to implement, easy to maintain, easy to understand, robust, and cheap (Kasabov 1996). The problem with applying fuzzy logic to filtration performance is that filtration performance is a complex process, and even though rule generation is easier than for expert systems, it may still be inadequate to describe the relationships involved in the process.

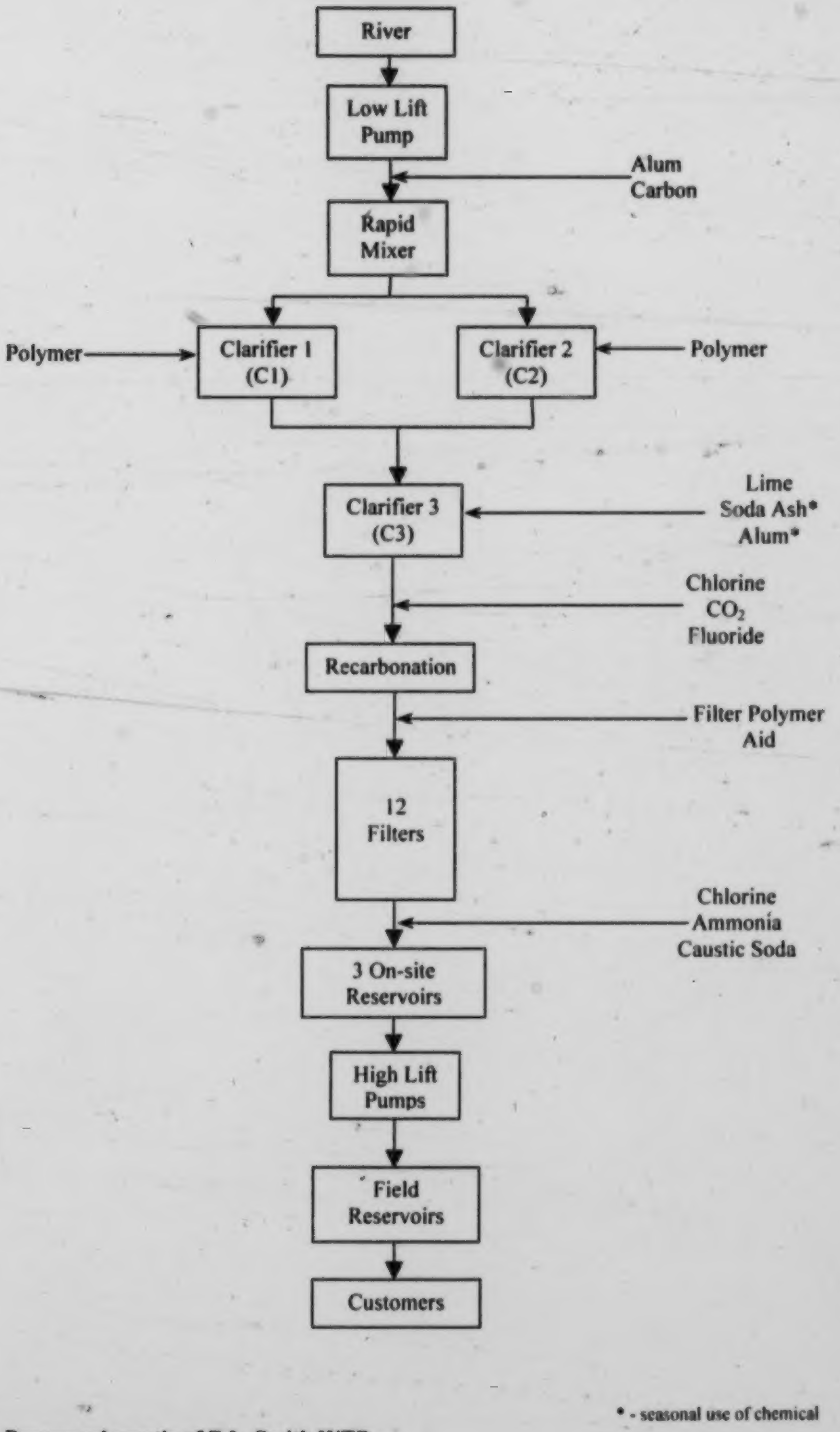
2.6.3 Hybrid Systems

There are systems in which rule-based systems, fuzzy logic, and neural networks, as well as other paradigms such as genetic algorithms and probabilistic reasoning, are combined together to form hybrid systems. These systems are used when requirements for the solution of a problem are not met with a single method, or if models either cannot be implemented using a single method or implement better using more than one method. Advantages of individual systems can thus be applied in a single system. If a fuzzy neural network were to be developed, then advantages of both systems could be utilised. For example, fuzzy systems can contribute the following features: a well-developed fuzzy logic theory, humanlike reasoning mechanisms, uses linguistic terms, accommodates common-sense knowledge, ambiguous knowledge, imprecise but rational knowledge, uses universal approximation techniques, robustness, fault-tolerance, low cost of development and maintenance, and low computational cost. Neural networks can contribute the following advantages: ability to learn from data, ability to model empirical behaviour of humans, use of universal approximation techniques, good generalisation, can extract knowledge from data, possess methods for data analysis, associative memories and pattern-matching techniques, massive parallelism (during data processing, many neurons may "fire" simultaneously), and robustness (Kasabov 1996). There are also several disadvantages of using hybrid systems. Knowledge and skills about different paradigms and methods are required, investigation and comparison of alternative solutions is necessary, and suitable software environment and simulation equipment is needed (Kasabov 1996).

2.7 E.L. Smith WTP Description

The E.L. Smith WTP was built in 1976 and is owned and operated by EPCOR. It is located on the western side of the city of Edmonton, AB, and it draws water from the North Saskatchewan River, which is fed by the Saskatchewan Glacier on the Columbia Icefields. The plant was expanded in 1984 and it currently has a design capacity of approximately 200 ML/d, but it can be further expanded to a capacity of 800 ML/d.

The first stage in water treatment involves adding alum for coagulation / flocculation and powdered activated carbon (PAC) for colour removal and taste and odour control to water pumped in from the North Saskatchewan River upstream of the clarifiers. Rapid mixing occurs by means of an inline mixer. Raw water then enters one of two upflow solids-contacting clarifiers running parallel. Polymer is introduced to the clarifiers to assist in flocculation. Entry of the influent raw water is through a circular draught tube with an impeller mixer in the clarifier basin, and recirculated sludge is mixed in with the raw water. Water is forced upward and out of the draught tube and into the basin because of the recirculated flow generated by the mixer. Water may then re-enter the draught tube through the bottom or exit the clarifier through tube settlers. After leaving the clarifiers, lime is added for softening (soda ash is also added seasonally for the same purpose) to the reaction zone of the third clarifier. After exiting the third clarifier, pH is adjusted by the addition of CO_2, and chlorine is added for disinfection as well as for further taste and odour control. Fluoride is added at this point as well. Water then flows by gravity through one of twelve filters in parallel with some polymer added as a filter aid. The filters are dual media rapid sand filters consisting of 475 mm of anthracite supported on 300 mm of sand and are approximately 18.9 m by 6.7 m length by width. After leaving the filters, chlorine and ammonia (for disinfection through the distribution system) and caustic soda (for pH adjustment) is added to the combined filter effluent before being pumped out to the reservoirs and into the distribution system. A schematic of the treatment process at E.L. Smith WTP is shown in Figure 2-6.



River
Low Lift Pump
Alum
Carbon
Rapid Mixer
Polymer
Clarifier 1 (C1)
Clarifier 2 (C2)
Polymer
Clarifier 3 (C3)
Lime
Soda Ash*
Alum*
Chlorine
CO_2
Fluoride
Recarbonation
Filter Polymer Aid
12 Filters
Chlorine
Ammonia
Caustic Soda
3 On-site Reservoirs
High Lift Pumps
Field Reservoirs
Customers
* - seasonal use of chemical

Figure 2-6. Process schematic of E.L. Smith WTP.

3.0 Methodology

3.1 Introduction

Described in this section are the steps and tasks that were completed in order to develop an ANN model of filtration performance. There were three main steps involved in creating a model: 1) source data analysis, 2) data preparation, and 3) model development and optimisation.

Source data analysis involves the evaluation of the available data in terms of what data are available and determining the inputs and outputs to be used in the model based on availability and significance as discussed in section 2.4. Also part of source data analysis was the time scale analysis to determine which time scale of data readings was most appropriate whether every minute, an hourly average, or a daily average. Characteristics of the data set of the chosen factors were also examined.

Data preparation involves setting up the data patterns in preparation for model development and optimisation, with data patterns defined as the set of input values and output value(s) for the one day. Part of the preparation of the data is to divide the patterns into a training set, a testing set, and a production set as explained in section 3.3.4. Normally, model development would commence after this point. However, preliminary attempts at development were unsuccessful due in part to large variations within the data making it difficult for the models to converge and generalise. The data was therefore further processed to aid in model development. Part of the processing was the categorisation of the data patterns into a number of sets. The purpose of this was to address the problems of large variations in the data by grouping together data patterns that shared similar characteristics. A model could then be developed for each category. Another problem contributing to poor model development was the presence of noise patterns. Noise is defined here as data patterns with similar input values, but greatly different output values, and this results in difficulties in model convergence and generalisation. The occurrence of noise can arise from poor readings due to instrumentation error. It can also arise from operational activities at the WTP in which abnormal operations are occurring on particular days such as the shutdown of a clarifier or filter. Therefore, such data patterns can be removed since they do not represent the normal operations of the WTP, which is what the model represents. Consequently, the removal of such noise patterns should result in improved model development.

Finally, with the data set prepared, model development and optimisation can take place. In this case, models of each category were developed. This involved testing various network architectures and numerous network parameters, such as number of neurons and type of activation function, until the best model performance is achieved for each model. The final overall model configuration was also examined to determine how each model would work together as an overall predictor of filtration performance.

3.2 Source Data Analysis

One of the initial tasks to perform prior to modelling is the analysis of the data set to be used for modelling. This includes identifying what data are available, determining what inputs to include in the model, and analysing the appropriateness of the data for modelling. It is also important to determine how much data are available.

3.2.1 Data Availability

Table 3-1 lists data that are recorded and available at the E.L. Smith WTP. At the time of model development, data from March 14, 1998 to April 30, 1999 was utilised.

3.2.2 Input and Output Selection

For the output, reservoir influent particle counts (>2 µm) was chosen, and this is based on the fact that particle counts are more sensitive to changes in the treatment process compared to turbidity as explained in section 2.2. The size range of >2 µm was chosen over the size range of 2-15 µm since the >2 µm size range essentially is a measure of total counts, whereas 2-15 µm is a more limited range in comparison. Therefore, the size range of >2 µm is a more accurate measure of finished water quality. Filter effluent particle counts of each filter was not chosen as outputs for a number of reasons. One reason is that it makes the configuration of the model unnecessarily complicated in using twelve outputs compared to using just one output, or alternatively developing twelve models for each filter compared to just one model. More importantly is the idea behind developing this model, which is to generate a model that encompasses all unit processes within water treatment rather than isolating filtration by itself and modelling that process.

Table 3-1. Available data at E.L. Smith WTP.

Description of Data	Data Type
Raw water particle counts (2-15 μm)	Water characteristics
Raw water particle counts (>2 μm)	Water characteristics
Recarbonation influent particle counts (2-15 μm)	Water characteristics
Recarbonation influent particle counts (>2 μm)	Water characteristics
Filters 1-12 effluent particle counts (2-15 μm)	Performance parameter
Filters 1-12 effluent particle counts (>2 μm)	Performance parameter
Clarifier 3 effluent particle counts (2-15 μm)	Water characteristics
Clarifier 3 effluent particle counts (>2 μm)	Water characteristics
Reservoir influent particle counts (2-15 μm)	Performance parameter
Reservoir influent particle counts (>2 μm)	Performance parameter
Raw water turbidity (daily high)	Water characteristics
Raw water turbidity (bench test)	Water characteristics
Raw water turbidity (online)	Water characteristics
Recarbonation influent turbidity	Water characteristics
Filters 1-12 effluent turbidity – (onstream)	Performance parameter
Filters 1-12 effluent turbidity – (bench test)	Performance parameter
Clarifier 3 effluent turbidity	Water characteristics
Reservoir influent turbidity	Performance parameter
Raw water temperature	Water characteristics
Raw water colour (daily high)	Water characteristics
Raw water colour (daily average)	Water characteristics
Total hardness	Water characteristics
Total alkalinity	Water characteristics
Maximum conductivity	Water characteristics
Clarifier 3 influent pH	Water characteristics
Clarifier 3 effluent pH	Water characteristics
Recarbonation influent pH	Water characteristics
Recarbonation effluent pH	Water characteristics
Filter influent pH	Water characteristics
Alum dose	Operational data
Lime dose	Operational data
Chlorine dose	Operational data
PAC dose	Operational data
Ammonia dose	Operational data
Soda ash dose	Operational data
Primary polymer dose	Operational data
Raw water flow	Operational data
Filter 1-12 effluent flow	Operational data
Noon air temperature	Miscellaneous

In taking this approach, reservoir influent particle counts is a better measure for taking into account the overall effectiveness of water treatment rather than each filter individually. Moreover, it is the finished water quality that proceeds out into the distribution system that is of greater importance than the effluent of each filter. In terms of operations, changes in operations upstream generally are performed in reaction to overall finished water quality rather than individual filter performance, and this is another reason for using reservoir influent particle counts rather than individual filter effluent particle counts.

The following inputs were chosen to be included in the model: raw water turbidity (bench test), raw water temperature, total hardness, total alkalinity, recarbonation effluent pH, alum dose, lime dose, PAC dose, polymer dose, raw flow (total flow through the plant). As discussed in section 2.4, all of these parameters have a significant impact on filtration performance, and since all of these readings were available, these ten parameters were chosen on this basis. The appendix section on pages 140-144 describes how they were measured and the related protocols for measurement.

The other data were not included as inputs for several reasons. Filter media size and depth are considered to be important factors in influencing filtration performance. However, since these factors are fixed in an operational water treatment plant, they cannot be changed nor controlled and therefore is not suitable to be included in a neural network model. The type of coagulant and polymer is also not a significant factor since it does not change during operation at E.L. Smith WTP, therefore they were not included as inputs. However, this type of parameter would be included in WTPs that change coagulants or polymers during the year. Ionic strength, particle size and size distribution were not included as inputs because they are not measured. Maximum conductivity has an influence on filtration performance, however the accuracy of the instrumentation readings is questionable, thus it was not included as an input. As a measure of levels of particulate matter entering the plant, turbidity was chosen over particle counts. The main reason for this is that at high levels of particulate matter, turbidity is a fairly accurate measure. Particle counts are less accurate in this case in that the particle counters themselves are limited in accuracy at high levels of particulate matter. Therefore, particle counts was not included as an input. Clarifier 3 (C3) effluent turbidity could have been a potential input rather than raw turbidity. However, preliminary model runs indicated better model performance with raw turbidity rather than C3 effluent turbidity, therefore it was decided to stay with raw turbidity. Little information regarding the impact of chlorine dose and ammonia dose on filtration performance was encountered in the literature, therefore they were not included as inputs. Raw colour, which is related to organic matter concentration, was also not included as an input. Since PAC dose, which impacts colour and organic matter concentration, is already included as an input, it would be redundant to include colour as an input as well as PAC dose. Soda ash dose was not included because it was rarely used during treatment, therefore it was unnecessary to include it. Noon air temperature was not included since raw water temperature has a direct impact on filtration performance rather than air temperature. Effluent flow from filters 1-12 was not included as inputs since each individual filter was not being modelled.

pH readings are available at various points in the treatment process, thus choosing between the different sets of pH readings was not straightforward. As discussed in section 2.4.2.3, pH impacts filtration performance by influencing the solubility of alum for coagulation and by affecting the surface charge of particles. However, since all pH readings that are available are post coagulation, the effect of pH on the particles themselves should be more greatly considered. Furthermore, such effects on particle surface charge are important because it will influence how the particle interacts with the filter media grains. Therefore, this effect would more likely be represented by pH readings nearer to filtration. With recarbonation for pH adjustment taking place, C3 influent and effluent pH and recarbonation influent pH is not as significant compared to recarbonation effluent pH and filter influent pH. Since filtration follows recarbonation, there should be little difference between the latter two pH readings. Some preliminary ANN models were therefore trained using one or the other to determine the impact each input had on prediction performance of the models. The result was that recarbonation effluent pH had a more significant positive impact in model prediction than filter influent pH, therefore recarbonation effluent pH was retained as an input.

3.2.3 Time Scale Analysis

Another factor to be considered is determining for what time scale the models should be developed. In other words, it should be considered if particle counts are to be predicted by the minute, as an hourly average, or as a daily average since data are available in each of these time scales. Developing a model to predict particle counts by the minute is impractical and unnecessary. Since water quality does not change drastically by the minute, a longer time scale is more relevant. The time scale up from minutes would be hourly average particle counts. However, a problem with hourly predictions is with the total retention time of water within the entire treatment process. The finished water that is measured for particle counts over an hour is not going to be the same as the raw water that is measured for turbidity and other raw water characteristics over that same hour. Thus, to compensate for this problem would require increasing the time scale, and daily averaged readings could accomplish this. The main concern is if particle count readings would be variable enough for ANN modelling to pick up the trends. In addition, it would need to be checked if trends in variability in hourly averaged readings are visible in daily averaged readings.

Figures 3-1 and 3-2 compare particle count readings by the minute, hourly averaged particle counts, and daily average particle counts. In Figure 3-1, which compares readings every minute versus the hourly average, it can be seen that the hourly average follows the definite trend displayed by the minute readings. The main difference is that there is more instrumentation noise associated with readings every minute. In Figure 3-2, which compares hourly average and daily average, it can also be seen that the daily average follows trends displayed by the hourly average. The main difference in this graph is that peaks are not as high for the daily average as for the hourly average. However, the daily average still demonstrates a fair amount of variability and therefore provides a sufficient description of the water quality. Furthermore, the peaks and trends that are observed in the daily average values are more likely to characterise the normal behaviour of treatment operations rather than hourly average values, in which peaks are more likely to be the result of upsets in the normal treatment process.

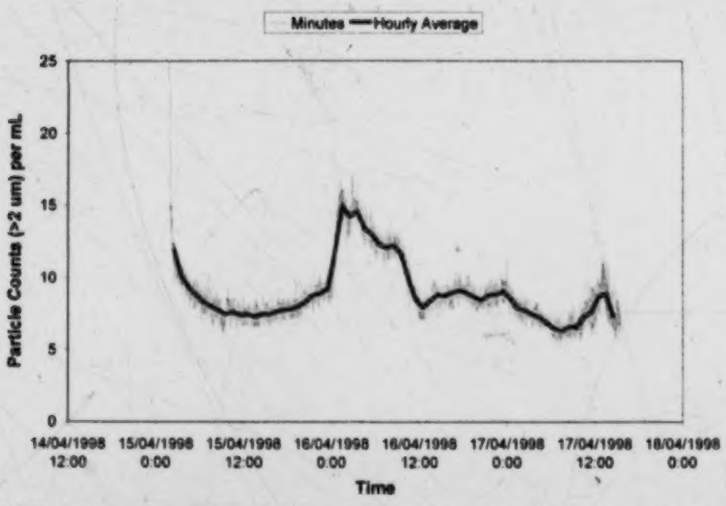


Figure 3-1. Minute readings vs. hourly average filter effluent particle counts from April 15-17, 1998 for filter 5 at E.L. Smith WTP.

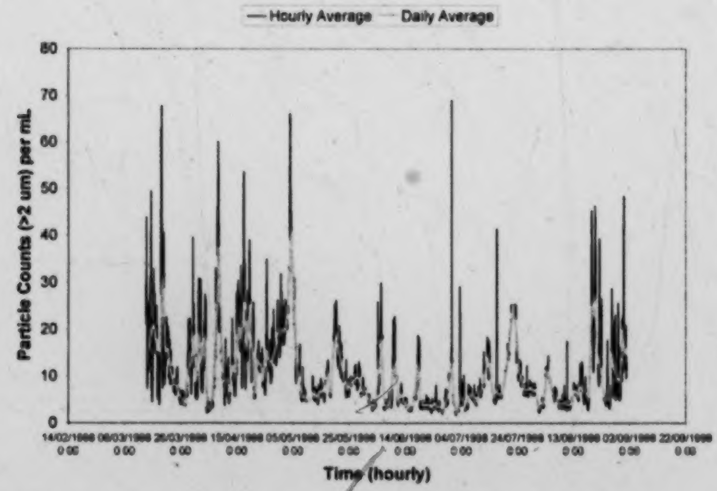


Figure 3-2. Hourly average vs. daily average reservoir influent particle counts from E.L. Smith WTP.

3.2.4 Data Characteristics

Table 3-2 lists some statistical information on each input and output that is used in model development.

Table 3-2. Statistics of input and output data to be used.

Input	Range	Average	Std. Dev.
raw turbidity (NTU)	2 - 1967	46.1	147.0
temperature (°C)	0.1 - 24.9	7.6	8.0
hardness (mg/L)	111 - 184	157.7	12.2
alkalinity (mg/L)	102 - 149	126.4	7.1
alum dose (mg/L)	18 - 198	44.9	29.2
lime dose (mg/L)	38 - 93	57.6	10.9
PAC dose (mg/L)	0 - 158	7.0	22.8
polymer dose (mg/L)	0.11 - 0.43	0.23	0.08
raw flow (ML/d)	81 - 253	187.9	22.4
recarb. effluent pH	7.61 - 8.37	7.96	0.16
reservoir influent particle counts (counts/mL)	1.5 - 115.3	12.3	11.6

As it can be seen, there is much variability in raw water quality throughout an entire year, and in response, operational parameters can vary just as much. With raw water turbidity varying between 2-1967 NTU and temperature varying between 0.1-24.9°C, it is clear that variability is very high. These are not the sorts of conditions that are normally encountered for water treatment in other regions, and therefore it is difficult to model under such conditions. Finished water quality, however, tends to remain at a constant level rarely exceeding 50 particle counts/mL.

3.3 Methodology of Data Categorisation and Preparation

With the available data analysed and the inputs and output chosen, the next stage is to prepare the data for model development. As previously mentioned, data patterns would normally be divided into a training set, a testing set, and a production set followed by model development. However, initial attempts at model development were unsuccessful, therefore the data patterns needed to be further processed. As part of the processing, the data patterns are to be categorised to group together similar patterns to deal with large variations thereby assisting in model convergence and generalisation. This procedure is outlined in this section.

A total of 410 data patterns were available starting from March 14, 1998 up to April 30, 1999. Data were categorised in two stages using a Kohonen network followed by a noise analysis. An explanation of how the network performs the categorisation is given in section 2.5.4, and it is used because it allows for classification based on multiple inputs as opposed to graphical techniques that only allow for three inputs in three dimensions.

3.3.1 First Stage Categorisation

The first categorisation was based on raw water data inputs of raw turbidity, temperature, total hardness, and total alkalinity, and the result was two categories of data patterns. Of the 410 data patterns, 221 or 53.9% of the data patterns went into category 1, 138 or 33.7 % of the data patterns went into category 2, and 51 or 12.4% did not go strictly into either category. Table 3-3 shows the values of each category, and it can be seen that main differences between the categories are raw water turbidity and temperature.

Table 3-3. Statistical values of input data in each category.

	category 1			category 2		
Inputs	range	average	std dev	range	average	std dev
raw turbidity (NTU)	2 - 841	22.2	76.8	2 - 720	56.8	107.0
temperature (°C)	0.1 - 9.5	1.5	2.1	9.1 - 24.9	16.8	3.7
hardness (mg/L)	126 - 184	161.3	12.2	140 - 181	153.6	8.0
alkalininty (mg/L)	108 - 144	126.8	7.2	111 - 143	126.2	5.7

*calculated after placement of the non-specific data

Of the 51 data patterns that were not categorised by the Kohonen network, 44 patterns could be placed in either category 1 or 2, while the remaining 7 were extreme values that could not go into either category. In Figure 3-3 below, category 1 data are shown as circles, category 2 data are shown as squares, extreme values are shown as triangles, and patterns which can be placed in either category are shown as an X. The distinction between the two categories is clearly illustrated to be based on temperature.

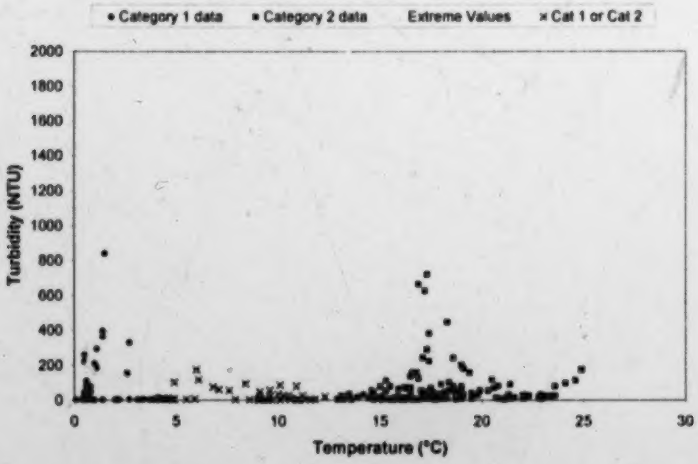


Figure 3-3. Distinction between category 1 and 2 based on turbidity and temperature.

The 44 data patterns that could go into either category were placed according to whichever category the pattern was closest to as determined by the Kohonen network. As a result, 19 of these patterns were placed in category 1 for a total of 240 patterns in that category, and 25 were placed in category 2 for a total of 163 patterns in that category. This resulted in the definition of the categories shown in Figure 3-4.

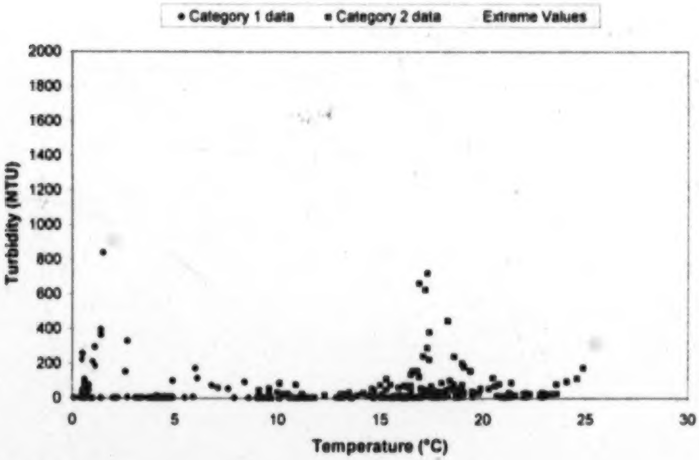


Figure 3-4. Distinction between category 1 and 2 based on turbidity and temperature after manual placement.

Although there is slight overlap between the two categories, the boundaries are still fairly well defined.

The remaining seven extreme data patterns include:

Date	Description
July 1-2, 1998	Extremely high raw turbidity values
April 12-16, 1999	Low alkalinity and hardness

These patterns have input values beyond the domain of the two defined categories. Figure 3-5 graphically shows how far out their values are compared to the categories themselves.

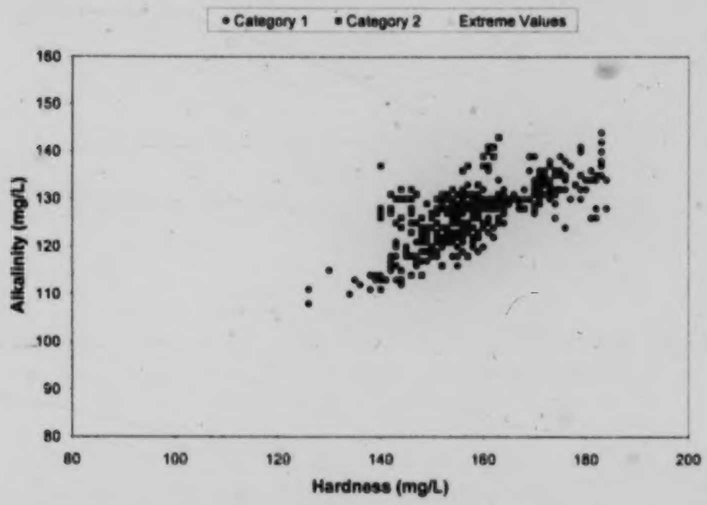


Figure 3-5. Comparison of extreme data points to category 1 and 2 based on alkalinity and hardness.

2

Table 3-4 shows the input values of the extremities, and it clearly shows that they are indeed extreme cases when compared to the average values of each category.

Table 3-4. Values of extreme cases compared to category average.

Date	raw turbidity (NTU)	temperature (°C)	hardness (mg/L)	alkalininty (mg/L)
01-Jul-98	1967	16	158	149
02-Jul-98	1306	16.7	143	146
12-Apr-99	207	0.6	120	105
13-Apr-99	304	0.6	111	102
14-Apr-99	187	0.9	113	107
15-Apr-99	187	0.9	113	107
16-Apr-99	162	1	119	107
category 1 avg.	22.2	1.5	161.3	126.8
category 2 avg.	56.8	16.8	153.6	126.2

Since these patterns do not represent the typical scenarios encountered in water treatment, these patterns would hinder model development. Therefore, these patterns were excluded from the data set at this point in model development. However, these patterns may be incorporated into the model after further development.

3.3.2 Second Stage Categorisation

The second stage categorisation was based on the operational parameters of alum dosage, lime dosage, PAC dosage, polymer dosage, raw flow, and the water characteristic parameter of recarbonation effluent pH. This categorisation was done to each of the two categories that resulted from first stage categorisation, and this further processing was done to further assist in preparing data for data set extraction, which will be discussed in section 3.3.4.

3.3.2.1 Category 1

Out of 240 patterns, 75 or 31.3% went into sub-category 1, 75 or 31.3% went into sub-category 2, and 90 or 37.5% were not placed in either sub-category. The reasons for non-placement are the same as in the first stage. Either it was unclear as to which sub-category the pattern solely belonged to in that it could go into either sub-category, or the pattern did not fit into either sub-category at all.

72 of the 90 uncategorised data were manually placed into one or the other sub-category based on whichever sub-category the pattern was closest to as calculated by the Kohonen network. 47 patterns went to sub-category 1 for a total of 122 patterns, and 25 patterns

went into sub-category 2 for a total of 100 patterns. Table 3-5 shows the values of the inputs of each sub-category.

Table 3-5. Statistical values of input data in each sub-category of category 1.

	sub-category 1			sub-category 2		
Inputs	range	average	std. dev.	range	average	std. dev.
alum dose (mg/L)	18 - 77	34.7	10.5	31 - 39	34.3	1.3
lime dose (mg/L)	43 - 68	53.0	6.4	52 - 77	62.7	5.4
PAC dose (mg/L)	0 - 44.6	7.6	12.2	0 - 3.2	0.2	0.7
polymer dose (mg/L)	0.13 - 0.36	0.23	0.04	0.26 - 0.43	0.33	0.04
raw flow (ML/d)	149 - 222	190.4	16.4	156 - 199	169.9	9.9
pH	7.66 - 8.15	7.93	0.12	7.61 - 8.05	7.88	0.12

The main differences seem to be a lower average lime dose, higher PAC dose, and higher raw flow in sub-category 1 than in sub-category 2. As well, alum dose in sub-category 1 is more scattered with a wider range than in sub-category 2, which is shown in Figure 3-6. Sub-category 1 is shown as circles, sub-category 2 is shown as squares, and the extreme cases are shown as triangles. Since only two dimensions are shown in the graph, it does not provide a complete illustration of how each sub-category is defined in n-dimensional space seeing as there are six inputs in total. However, the graph does show why the extreme values do not fit into either sub-category. The one exception that appears in the midst of sub-category 2 is an extreme value because of its low raw flow and high pH. Table 3-6 lists all 18 extreme cases with the extreme values highlighted, and when it is compared to the domain of each sub-category, it is clear that they do not fit into either category.

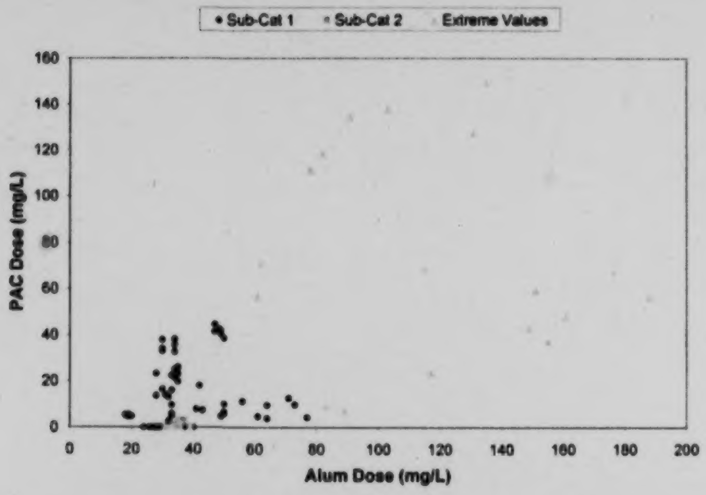


Figure 3-6. Comparison of extreme values to sub-category 1 and 2 of category 1 based on PAC dose and alum dose.

Table 3-6. Values of extreme cases compared to sub-category average.

Date	alum dose (mg/L)	lime dose (mg/L)	PAC dose (mg/L)	polymer dose (mg/L)	raw flow (ML/d)	pH
24-Feb-99	33	66	0	0.37	87	8.33
05-Apr-99	61	63	57	0.23	165	8.02
06-Apr-99	91	75	134.6	0.33	159	7.94
07-Apr-99	78	75	111.7	0.34	163	8.02
08-Apr-99	82	68	118.8	0.34	170	8.05
09-Apr-99	103	68	138	0.3	172	8.06
10-Apr-99	135	70	149.7	0.14	171	8.02
11-Apr-99	131	70	127.4	0.11	169	7.98
17-Apr-99	115	71	69	0.21	186	7.87
18-Apr-99	151	78	59.4	0.2	191	7.92
19-Apr-99	188	83	56.8	0.19	190	8.01
20-Apr-99	161	83	48.4	0.18	190	8.15
21-Apr-99	149	84	43	0.27	189	8.19
22-Apr-99	155	85	37.4	0.2	190	8.05
23-Apr-99	117	81	23.7	0.14	181	7.91
24-Apr-99	92	77	22.1	0.14	181	7.64
29-Apr-99	83	65	9.5	0.12	190	7.87
30-Apr-99	89	68	7.4	0.12	190	7.75
sub-cat 1 avg.	34.7	53.0	7.6	0.23	190.4	7.93
sub-cat 2 avg.	34.3	62.7	0.2	0.33	169.9	7.88

These values are not typical conditions encountered in water treatment within either sub-category, therefore they are excluded from further model development at this point to prevent hindrance of model development. They may, however, be used and incorporated into a model at a later point.

3.3.2.2 Category 2

Out of 163 patterns, 41 or 25.2% went into sub-category 1 and 54 or 33.1% went into sub-category 2. 68 or 41.7% were left uncategorised either because they could fit in both sub-categories or because they could not fit into either sub-category.

58 of the 68 uncategorised data patterns were manually placed into either sub-category depending on which sub-category was closest as calculated by the Kohonen network. 37 patterns were placed in sub-category 1 for a total of 78 patterns in that sub-category, and 21 patterns were placed in sub-category 2 for a total of 75 patterns. Table 3-7 shows the input values of each sub-category.

Table 3-7. Statistical values of input data in each sub-category of category 2.

	sub-category 1			sub-category 2		
Inputs	range	average	std. dev.	range	average	std. dev.
alum dose (mg/L)	26 - 124	51.7	24.8	23 - 80	35.4	14.3
lime dose (mg/L)	45 - 90	60.4	9.6	38 - 72	47.0	7.1
PAC dose (mg/L)	0 - 3.6	0.07	0.43	0 - 0	0.00	0.00
polymer dose (mg/L)	0.12 - 0.18	0.15	0.02	0.16 - 0.22	0.19	0.02
raw flow (ML/d)	100 - 235	195.8	18.6	170 - 253	206.6	24.8
pH	7.73 - 8.20	7.95	0.10	7.66 - 8.35	8.10	0.21

The most significant differences between the two sub-categories are that alum dose and lime dose is higher in sub-category 1 than in sub-category 2.

In Figure 3-7 showing the graph of alum dose vs. lime dose, sub-category 1 patterns are shown as circles, sub-category 2 patterns are shown as squares, and extreme cases are shown as triangles. Since this graph only shows two dimensions, it does not show a complete illustration defining each sub-category. However, even though there is some overlap between the sub-categories, which is due to the fact that the other four inputs are not represented in this graph, the extreme cases are clearly shown to stand out from the defined sub-categories. The four extreme cases that appear to be in the domain of the sub-categories are defined as extreme in that the raw flow and pH are beyond the domain in one case, and PAC dose and pH are beyond the sub-category domains in the other three cases. This is more clearly shown in Table 3-8 with the extreme input values highlighted.

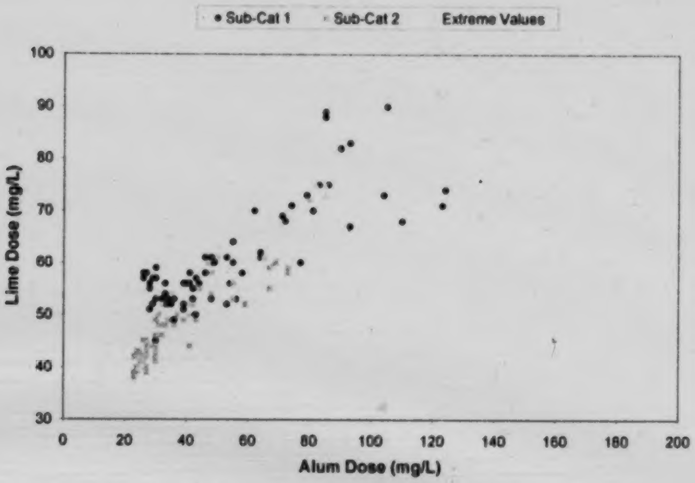


Figure 3-7. Comparison of extreme values to sub-category 1 and 2 of category 2 based on lime dose and alum dose.

Table 3-8. Values of extreme cases compared to sub-category average.

Date	alum dose (mg/L)	lime dose (mg/L)	PAC dose (mg/L)	polymer dose (mg/L)	raw flow (ML/d)	pH
03-Jul-98	172	91	0	0.13	170	8.19
04-Jul-98	143	80	0	0.12	170	8.19
07-Jul-98	137	84	0	0.16	172	8.19
08-Jul-98	124	93	0	0.16	179	8.20
09-Jul-98	122	92	0	0.16	179	8.25
10-Jul-98	121	93	0	0.16	179	8.19
16-Sep-98	25	40	0	0.18	81	8.37
25-Apr-99	85	75	22	0.14	190	7.65
26-Apr-99	85	73	20.3	0.14	189	7.61
27-Apr-99	82	75	17.9	0.14	190	7.62
sub-cat 1 avg.	51.7	60.4	0.07	0.15	195.8	7.95
sub-cat 2 avg.	35.4	47.0	0.00	0.19	206.6	8.10

Since these 10 patterns are extreme cases and are not typically encountered, they were excluded from the data set from further model development at this point. However, they may be incorporated into the model at a later point after models have been developed further.

3.3.3 Noise Analysis

As discussed in section 3.1, noise patterns are a problem that hinders model development. Noise is defined as data patterns with similar input values but vastly different output values. This would cause model predictions to be less accurate because there is conflict with the input-output cause-effect relationships the model is trying to capture. Therefore, with the removal of such patterns, convergence and generalisation should improve.

This analysis was performed on each sub-category separately and done in the following manner. A Kohonen network was used to categorise the data set into 20 different categories based on the 10 input parameters of raw turbidity, temperature, hardness, alkalinity, alum dose, lime dose, PAC dose, polymer dose, raw flow, and pH. This resulted in approximately 2-10 patterns in each category. The values of the input parameters of the patterns within a category should theoretically be quite similar. To determine if a pattern is considered to be noise, the value of the output parameter, reservoir influent particle count, was compared between the patterns within a category. The normalised value of the output was used to assist in evaluation with an explanation of its calculation given in section 3.3.4. If a data pattern had a significantly different value for its output parameter than the other patterns, then it was considered to be noise and is removed from the data set. All noise patterns were set aside to later be incorporated into a model depending on whether or not the data had an adverse effect on model predictions. Table 3-9 shows the total number of noise patterns that were removed from each category data set.

Table 3-9. Number of noise patterns removed from the category data sets.

Category	Number of total patterns	Number of noise patterns removed	% of total removed
cat 1 sub-cat 1	122	11	9.0%
cat 1 sub-cat 2	100	12	12.0%
cat 2 sub-cat 1	78	8	10.3%
cat 2 sub-cat 2	75	8	10.7%

The following pages provide an example of the noise analysis that was done for the cat 1 sub-cat 1 data set. Patterns removed are highlighted in grey.

Noise analysis example for cat 1 sub-cat 1

Category 1

DATE	Raw Turb Lab	TEMP (°C)	TOTAL HARD mg/L	TOTAL ALK. mg/L	Alum Plant #4	Lime Plant #4	Carbon Plant #4	Polymer Plant #4	Raw Flow	Recarb Effl pH	Res. Inf. particle counts (p/mL)
23-Mar-99	3	0.5	158	120	33	55	4.9	0.21	199	8.03	5.8
24-Mar-99	3	0.5	158	121	33	54	5.9	0.21	199	7.98	9.2
25-Mar-99	4	0.5	154	118	33	54	5.3	0.21	199	8.01	8.2
26-Mar-99	5	0.5	155	118	33	55	5.3	0.21	200	7.99	12.7
27-Mar-99	6	0.5	161	123	33	55	4.9	0.21	200	7.98	9.3
28-Mar-99	8	0.5	155	120	33	54	5.1	0.22	200	7.99	7.7
Normalised values											
23-Mar-99	0.00	0.04	0.59	0.36	0.25	0.63	0.11	0.42	0.62	0.77	0.12
24-Mar-99	0.00	0.04	0.59	0.40	0.25	0.58	0.14	0.42	0.62	0.65	0.21
25-Mar-99	0.01	0.04	0.49	0.28	0.25	0.58	0.12	0.42	0.62	0.71	0.19
26-Mar-99	0.01	0.04	0.51	0.28	0.25	0.63	0.12	0.42	0.64	0.67	0.31
27-Mar-99	0.02	0.04	0.67	0.48	0.25	0.63	0.11	0.42	0.64	0.65	0.22
28-Mar-99	0.03	0.04	0.51	0.36	0.25	0.58	0.12	0.47	0.64	0.68	0.17
Category 2											
29-Mar-99	8	0.5	155	116	42	53	18.1	0.22	185	8.00	7.5
Normalised values											
29-Mar-99	0.01	0.04	0.26	0.32	0.29	0.11	0.61	0.74	0.48	0.90	0.30
Category 3											
16-Mar-98	2	0.5	156	121	19	50	4.8	0.22	170	8.00	20.8
17-Mar-98	2	0.5	156	123	19	51	5.5	0.22	170	8.00	15.5
18-Mar-98	2	0.5	155	122	19	47	4.6	0.22	175	8.01	15.2
20-Mar-98	4	0.5	159	123	20	50	4.8	0.22	181	8.00	15.3
Normalised values											
16-Mar-98	0.00	0.04	0.54	0.40	0.02	0.37	0.11	0.47	0.15	0.70	0.54
17-Mar-98	0.00	0.04	0.54	0.48	0.02	0.42	0.13	0.47	0.15	0.70	0.39
18-Mar-98	0.00	0.04	0.51	0.44	0.02	0.21	0.11	0.47	0.23	0.72	0.38
20-Mar-98	0.01	0.04	0.62	0.48	0.03	0.37	0.11	0.47	0.33	0.69	0.38

Category 4

DATE	Raw Turb Lab	TEMP (°C)	TOTAL HARD mg/L	TOTAL ALK. mg/L	Alum Plant #4	Lime Plant #4	Carbon Plant #4	Polymer Plant #4	Raw Flow	Recarb Effl pH	Res. Inf. particle counts (p/mL)
14-Mar-98	2	0.5	163	126	19	53	4.8	0.22	167	8.11	15.8
15-Mar-98	2	0.5	160	125	18	51	5.3	0.22	161	8.03	20.0
Normalised values											
14-Mar-98	0.00	0.04	0.72	0.60	0.02	0.53	0.11	0.47	0.10	0.91	0.40
15-Mar-98	0.00	0.04	0.64	0.56	0.00	0.42	0.12	0.47	0.00	0.75	0.51
Category 5											
21-Mar-98	7	0.4	163	128	28	52	13.7	0.22	180	7.97	15.9
22-Mar-98	5	0.5	161	129	28	55	23.3	0.22	179	7.95	10.8
12-Nov-98	3	0.1	162	130	35	51	0	0.25	189	7.99	8.8
Normalised values											
21-Mar-98	0.02	0.03	0.72	0.68	0.17	0.47	0.32	0.47	0.31	0.64	0.40
22-Mar-98	0.01	0.04	0.67	0.72	0.17	0.63	0.54	0.47	0.30	0.60	0.26
12-Nov-98	0.00	0.00	0.69	0.76	0.29	0.42	0.00	0.63	0.46	0.68	0.20
Category 6											
17-Mar-99	3	0.5	163	128	32	58	0	0.19	171	8.09	1.5
18-Mar-99	6	0.5	163	126	31	57	0	0.19	175	8.09	9.6
19-Mar-99	4	0.5	171	134	31	56	14.1	0.19	188	8.09	7.6
Normalised values											
17-Mar-99	0.00	0.04	0.72	0.68	0.24	0.79	0.00	0.32	0.16	0.87	0.00
18-Mar-99	0.02	0.04	0.72	0.60	0.22	0.74	0.00	0.32	0.23	0.87	0.22
19-Mar-99	0.01	0.04	0.92	0.92	0.22	0.68	0.33	0.32	0.44	0.88	0.17

Category 7 DATE	Raw Turb Lab	TEMP (°C)	TOTAL HARD mg/L	TOTAL ALK. mg/L	Alum Plant #4	Lime Plant #4	Carbon Plant #4	Polymer Plant #4	Raw Flow	Recarb Effl pH	Res. Inf. particle counts (p/mL)
20-Mar-99	3	0.5	173	132	31	56	0	0.18	199	8.1	12.5
21-Mar-99	3	0.5	164	127	31	56	0	0.18	199	8.1	10.1
10-Mar-99	2	0.5	170	127	31	61	0	0.25	201	8.1	5.3
11-Mar-99	3	0.5	174	129	31	61	0	0.21	200	8.1	6.3
12-Mar-99	2	0.5	171	129	31	62	0	0.2	195	8.0	5.0
13-Mar-99	3	0.5	171	132	31	62	0	0.21	181	8.0	5.1
Normalised values											
20-Mar-99	0.00	0.04	0.97	0.84	0.22	0.68	0.00	0.26	0.62	0.86	0.30
21-Mar-99	0.00	0.04	0.74	0.64	0.22	0.68	0.00	0.26	0.62	0.85	0.24
10-Mar-99	0.00	0.04	0.90	0.64	0.22	0.95	0.00	0.63	0.66	0.80	0.10
11-Mar-99	0.00	0.04	1.00	0.72	0.22	0.95	0.00	0.42	0.64	0.81	0.13
12-Mar-99	0.00	0.04	0.92	0.72	0.22	1.00	0.00	0.37	0.56	0.76	0.10
13-Mar-99	0.00	0.04	0.92	0.84	0.22	1.00	0.00	0.42	0.33	0.67	0.10
Category 8											
06-Apr-98	15	0.5	143	113	32	47	13	0.32	180	8.00	6.1
Normalised values											
06-Apr-98	0.06	0.04	0.21	0.08	0.24	0.21	0.30	1.00	0.31	0.70	0.13

Category 9

DATE	Raw Turb Lab	TEMP (°C)	TOTAL HARD mg/L	TOTAL ALK. mg/L	Alum Plant #4	Lime Plant #4	Carbon Plant #4	Polymer Plant #4	Raw Flow	Recarb Effl pH	Res. Inf. particle counts (p/mL)
10-Oct-98	8	5.9	155	129	26	45	0	0.18	215	8.0	17.5
11-Oct-98	5	5.5	154	126	26	48	0	0.18	200	8.0	24.1
12-Oct-98	4	4.8	155	127	27	45	0	0.18	199	8.0	15.8
13-Oct-98	5	3.9	157	128	29	48	0	0.18	199	8.0	19.0
14-Oct-98	3	4.1	158	129	30	48	0	0.18	197	8.1	27.0
15-Oct-98	3	4.5	161	130	30	49	0	0.18	207	8.1	23.4
16-Oct-98	4	4.6	159	130	30	49	0	0.18	222	8.0	20.0
10-Oct-98	0.03	0.62	0.51	0.72	0.14	0.11	0.00	0.26	0.89	0.75	0.45
11-Oct-98	0.01	0.57	0.49	0.60	0.14	0.26	0.00	0.26	0.64	0.75	0.63
12-Oct-98	0.01	0.50	0.51	0.64	0.15	0.11	0.00	0.26	0.62	0.73	0.40
13-Oct-98	0.01	0.40	0.56	0.68	0.19	0.26	0.00	0.26	0.62	0.77	0.49
14-Oct-98	0.00	0.43	0.59	0.72	0.20	0.26	0.00	0.26	0.69	0.84	0.73
15-Oct-98	0.00	0.47	0.67	0.76	0.20	0.32	0.00	0.26	0.75	0.85	0.61
16-Oct-98	0.01	0.48	0.62	0.76	0.20	0.32	0.00	0.26	1.00	0.60	0.52

noise pattern marked in grey largely different from the lowest value (Oct. 12)

Category 10 DATE	Raw Turb Lab	TEMP (°C)	TOTAL HARD mg/L	TOTAL ALK. mg/L	Alum Plant #4	Lime Plant #4	Carbon Plant #4	Polymer Plant #4	Raw Flow	Recarb Effl pH	Res. Inf. particle counts (p/mL)
08-Oct-98	4	9.5	159	131	24	44	0	0.18	217	7.9	5.4
09-Oct-98	4	7.9	160	132	24	44	0	0.18	220	8.0	9.8
Normalised values											
08-Oct-98	0.01	1.00	0.62	0.80	0.10	0.05	0.00	0.26	0.92	0.42	0.11
09-Oct-98	0.01	0.83	0.64	0.84	0.10	0.05	0.00	0.26	0.97	0.73	0.23
Category 11											
20-Apr-98	54	7.6	157	123	49	58	4.9	0.16	189	7.7	15.9
02-Oct-98	2	9.1	161	136	24	49	0	0.2	179	7.7	25.6
03-Oct-98	4	9.3	158	130	24	48	0	0.21	189	7.7	25.5
04-Oct-98	3	9	160	128	24	47	0	0.22	211	7.8	25.5
05-Oct-98	3	8.6	159	131	24	46	0	0.22	212	7.8	25.3
06-Oct-98	2	9.1	157	129	24	48	0	0.21	211	7.8	18.0
07-Oct-98	2	9.5	158	127	24	45	0	0.19	211	7.8	5.3
Normalised values											
20-Apr-98	0.25	0.80	0.56	0.48	0.53	0.79	0.11	0.16	0.46	0.08	0.40
02-Oct-98	0.00	0.96	0.67	1.00	0.10	0.32	0.00	0.37	0.30	0.02	0.67
03-Oct-98	0.01	0.98	0.59	0.76	0.10	0.26	0.00	0.42	0.46	0.06	0.67
04-Oct-98	0.00	0.95	0.64	0.68	0.10	0.21	0.00	0.47	0.82	0.20	0.67
05-Oct-98	0.00	0.90	0.62	0.80	0.10	0.16	0.00	0.47	0.84	0.21	0.66
06-Oct-98	0.00	0.96	0.56	0.72	0.10	0.26	0.00	0.42	0.82	0.21	0.46
07-Oct-98	0.00	1.00	0.59	0.64	0.10	0.11	0.00	0.32	0.82	0.19	0.11

noise pattern marked in grey lower than the rest of the category

Category 12

DATE	Raw Turb Lab	TEMP (°C)	TOTAL HARD mg/L	TOTAL ALK. mg/L	Alum Plant #4	Lime Plant #4	Carbon Plant #4	Polymer Plant #4	Raw Flow	Recarb Effl pH	Res. Inf. particle counts (p/mL)
17-Oct-98	11	4.2	164	130	31	52	0	0.2	220	7.8	29.3
25-Oct-98	4	4.3	157	133	31	51	0	0.26	221	7.8	13.7
26-Oct-98	4	4.4	156	129	30	51	0	0.25	220	7.8	13.3
27-Oct-98	5	4.5	157	129	30	50	0	0.25	220	7.8	13.5
28-Oct-98	5	4.8	158	131	30	49	0	0.25	221	7.8	15.5
29-Oct-98	4	3.9	155	128	30	49	0	0.25	209	7.7	9.3
Normalised values											
17-Oct-98	0.04	0.44	0.74	0.76	0.22	0.47	0.00	0.37	0.97	0.25	0.77
25-Oct-98	0.01	0.45	0.56	0.88	0.22	0.42	0.00	0.68	0.98	0.28	0.34
26-Oct-98	0.01	0.46	0.54	0.72	0.20	0.42	0.00	0.63	0.97	0.29	0.33
27-Oct-98	0.01	0.47	0.56	0.72	0.20	0.37	0.00	0.63	0.97	0.35	0.33
28-Oct-98	0.01	0.50	0.59	0.80	0.20	0.32	0.00	0.63	0.98	0.23	0.39
29-Oct-98	0.01	0.40	0.51	0.68	0.20	0.32	0.00	0.63	0.79	0.16	0.22

noise pattern marked in grey higher than rest of category

Category 13

DATE	Raw Turb Lab	TEMP (°C)	TOTAL HARD mg/L	TOTAL ALK. mg/L	Alum Plant #4	Lime Plant #4	Carbon Plant #4	Polymer Plant #4	Raw Flow	Recarb Effl pH	Res. Inf. particle counts (p/mL)
18-Oct-98	15	4.1	159	129	35	54	0	0.22	209	7.8	17.2
19-Oct-98	11	4.2	156	128	37	54	0	0.25	207	7.8	18.3
20-Oct-98	10	4.5	152	129	40	55	0	0.29	204	7.9	17.9
21-Oct-98	8	4.7	156	129	40	55	0	0.29	217	7.9	12.5
22-Oct-98	7	4.9	155	129	39	54	0	0.29	220	7.9	9.6
23-Oct-98	5	4.8	153	130	40	54	0	0.28	220	7.8	9.1
24-Oct-98	6	4.4	156	131	36	53	0	0.28	221	7.8	8.6
Normalised values											
18-Oct-98	0.06	0.43	0.62	0.72	0.29	0.58	0.00	0.47	0.79	0.37	0.44
19-Oct-98	0.04	0.44	0.54	0.68	0.32	0.58	0.00	0.63	0.75	0.36	0.47
20-Oct-98	0.04	0.47	0.44	0.72	0.37	0.63	0.00	0.84	0.70	0.43	0.46
21-Oct-98	0.03	0.49	0.54	0.72	0.37	0.63	0.00	0.84	0.92	0.56	0.30
22-Oct-98	0.02	0.51	0.51	0.72	0.36	0.58	0.00	0.84	0.97	0.48	0.22
23-Oct-98	0.01	0.50	0.46	0.76	0.37	0.58	0.00	0.79	0.97	0.29	0.21
24-Oct-98	0.02	0.46	0.54	0.80	0.31	0.53	0.00	0.79	0.98	0.29	0.20

Category 14

DATE	Raw Turb Lab	TEMP (°C)	TOTAL HARD mg/L	TOTAL ALK. mg/L	Alum Plant #4	Lime Plant #4	Carbon Plant #4	Polymer Plant #4	Raw Flow	Recarb Effl pH	Res. Inf. particle counts (p/mL)
15-Nov-98	3	0.9	161	126	32	51	0	0.25	199	7.80	14.9
16-Nov-98	7	0.7	162	128	35	51	0	0.25	199	7.79	12.7
17-Nov-98	10	0.5	164	132	34	53	0	0.25	199	7.81	6.0
18-Nov-98	14	0.5	167	129	35	56	0	0.24	199	7.81	12.9
21-Nov-98	5	0.5	166	131	34	58	0	0.25	199	7.81	11.5
22-Nov-98	5	0.5	163	129	30	57	0	0.26	199	7.85	7.8
23-Nov-98	3	0.5	161	128	32	55	0	0.26	190	7.86	7.2
10-Nov-98	3	0.4	156	128	35	54	0	0.25	181	7.75	28.6
11-Nov-98	3	0.1	158	126	35	52	0	0.25	184	7.89	21.2
13-Nov-98	3	0.6	163	134	35	51	0	0.25	192	7.88	10.2
Normalised values											
15-Nov-98	0.00	0.09	0.67	0.60	0.24	0.42	0.00	0.63	0.62	0.29	0.37
16-Nov-98	0.02	0.06	0.69	0.68	0.29	0.42	0.00	0.63	0.62	0.27	0.31
17-Nov-98	0.04	0.04	0.74	0.84	0.27	0.53	0.00	0.63	0.62	0.32	0.12
18-Nov-98	0.06	0.04	0.82	0.72	0.29	0.68	0.00	0.58	0.62	0.31	0.32
21-Nov-98	0.01	0.04	0.79	0.80	0.27	0.79	0.00	0.63	0.62	0.31	0.28
22-Nov-98	0.01	0.04	0.72	0.72	0.20	0.74	0.00	0.68	0.62	0.38	0.17
23-Nov-98	0.00	0.04	0.67	0.68	0.24	0.63	0.00	0.68	0.48	0.41	0.16
10-Nov-98	0.00	0.03	0.54	0.68	0.29	0.58	0.00	0.63	0.33	0.19	0.75
11-Nov-98	0.00	0.00	0.59	0.60	0.29	0.47	0.00	0.63	0.38	0.48	0.85
13-Nov-98	0.00	0.05	0.72	0.92	0.29	0.42	0.00	0.63	0.51	0.45	0.24

noise patterns marked in grey higher than rest of category

Category 15

DATE	Raw Turb Lab	TEMP (°C)	TOTAL HARD mg/L	TOTAL ALK. mg/L	Alum Plant #4	Lime Plant #4	Carbon Plant #4	Polymer Plant #4	Raw Flow	Recarb Effl pH	Res. Inf. particle counts (p/mL)
30-Oct-98	6	3.4	154	125	29	46	0	0.25	201	7.68	18.7
06-Nov-98	3	3.3	162	124	30	48	0	0.23	181	7.85	19.4
07-Nov-98	3	2	159	129	30	47	0	0.24	180	7.84	21.8
08-Nov-98	3	1.4	160	130	30	47	0	0.24	180	7.81	23.1
09-Nov-98	2	1.4	160	130	31	49	0	0.24	185	7.78	18.1
14-Nov-98	3	0.6	158	129	31	48	0	0.25	197	7.80	17.3
Normalised values											
30-Oct-98	0.02	0.35	0.49	0.56	0.19	0.16	0.00	0.63	0.66	0.05	0.48
06-Nov-98	0.00	0.34	0.69	0.52	0.20	0.26	0.00	0.53	0.33	0.39	0.50
07-Nov-98	0.00	0.20	0.62	0.72	0.20	0.21	0.00	0.58	0.31	0.38	0.56
08-Nov-98	0.00	0.14	0.64	0.76	0.20	0.21	0.00	0.58	0.31	0.31	0.60
09-Nov-98	0.00	0.14	0.64	0.76	0.22	0.32	0.00	0.58	0.39	0.25	0.46
14-Nov-98	0.00	0.05	0.59	0.72	0.22	0.26	0.00	0.63	0.59	0.30	0.44

Category 16

DATE	Raw Turb Lab	TEMP (°C)	TOTAL HARD mg/L	TOTAL ALK. mg/L	Alum Plant #4	Lime Plant #4	Carbon Plant #4	Polymer Plant #4	Raw Flow	Recarb Effl pH	Res. Inf. particle counts (p/mL)
31-Oct-98	6	2.7	157	124	30	45	0	0.25	201	7.79	22.2
01-Nov-98	5	2.2	156	123	30	46	0	0.23	201	7.84	16.4
02-Nov-98	4	2.1	154	123	30	45	0	0.23	200	7.83	17.7
03-Nov-98	4	2.1	151	119	29	44	0	0.23	200	7.80	19.9
04-Nov-98	4	3.1	151	118	28	43	0	0.23	197	7.81	19.0
05-Nov-98	3	3.7	155	120	29	46	0	0.23	184	7.85	15.8
Normalised values											
31-Oct-98	0.02	0.28	0.56	0.52	0.20	0.11	0.00	0.63	0.66	0.26	0.58
01-Nov-98	0.01	0.22	0.54	0.48	0.20	0.16	0.00	0.53	0.66	0.37	0.42
02-Nov-98	0.01	0.21	0.49	0.48	0.20	0.11	0.00	0.53	0.64	0.36	0.45
03-Nov-98	0.01	0.21	0.41	0.32	0.19	0.05	0.00	0.53	0.64	0.29	0.51
04-Nov-98	0.01	0.32	0.41	0.28	0.17	0.00	0.00	0.53	0.59	0.31	0.49
05-Nov-98	0.00	0.38	0.51	0.36	0.19	0.16	0.00	0.53	0.38	0.39	0.40

Category 17

DATE	Raw Turb Lab	TEMP (°C)	TOTAL HARD mg/L	TOTAL ALK. mg/L	Alum Plant #4	Lime Plant #4	Carbon Plant #4	Polymer Plant #4	Raw Flow	Recarb Effl pH	Res. Inf. particle counts (p/mL)
30-Mar-98	4	0.5	147	116	35	45	25.6	0.28	189	8.01	18.5
31-Mar-98	5	0.5	149	117	33	46	22.4	0.3	186	8.00	8.9
01-Apr-98	4	0.5	148	117	35	45	23	0.29	180	8.00	18.9
02-Apr-98	4	0.5	143	121	34	44	21.5	0.29	179	8.00	13.0
03-Apr-98	5	0.5	142	118	35	44	19.8	0.29	180	8.00	18.1
04-Apr-98	7	0.5	144	112	30	45	16.5	0.29	180	7.99	3.7
05-Apr-98	11	0.5	144	112	33	43	16	0.3	173	8.00	3.3
07-Apr-98	21	0.5	136	112	33	44	9.6	0.28	177	8.00	17.4
08-Apr-98	28	0.5	135	113	41	52	7.9	0.28	179	8.00	37.4
Normalised values											
30-Mar-98	0.01	0.04	0.31	0.20	0.29	0.11	0.60	0.79	0.46	0.73	0.47
31-Mar-98	0.01	0.04	0.36	0.24	0.25	0.16	0.52	0.89	0.41	0.70	0.21
01-Apr-98	0.01	0.04	0.33	0.24	0.29	0.11	0.54	0.84	0.31	0.69	0.48
02-Apr-98	0.01	0.04	0.21	0.40	0.27	0.05	0.50	0.84	0.30	0.70	0.32
03-Apr-98	0.01	0.04	0.18	0.28	0.29	0.05	0.46	0.84	0.31	0.69	0.46
04-Apr-98	0.02	0.04	0.23	0.04	0.20	0.11	0.39	0.84	0.31	0.69	0.06
05-Apr-98	0.04	0.04	0.23	0.04	0.25	0.00	0.37	0.89	0.20	0.70	0.05
07-Apr-98	0.09	0.04	0.03	0.04	0.25	0.05	0.22	0.79	0.26	0.70	0.44
08-Apr-98	0.13	0.04	0.00	0.08	0.39	0.47	0.18	0.79	0.30	0.69	1.00

noise patterns marked in grey higher and lower than the rest of category

Category 18

DATE	Raw Turb Lab	TEMP (°C)	TOTAL HARD mg/L	TOTAL ALK. mg/L	Alum Plant #4	Lime Plant #4	Carbon Plant #4	Polymer Plant #4	Raw Flow	Recarb Effl pH	Res. Inf. particle counts (p/mL)
23-Mar-98	4	0.5	157	124	30	53	37.9	0.27	182	8.10	8.2
24-Mar-98	3	0.5	152	120	30	51	34	0.27	187	8.15	8.0
25-Mar-98	4	0.5	149	117	34	46	38.2	0.27	182	8.12	6.1
26-Mar-98	4	0.5	150	119	34	48	35.7	0.29	181	8.11	4.8
27-Mar-98	5	0.5	150	119	34	49	32.7	0.29	181	8.09	4.2
28-Mar-98	4	0.5	149	122	30	47	33.1	0.29	182	8.02	11.1
29-Mar-98	5	0.5	145	119	35	45	26	0.27	190	8.10	12.4
Normalised values											
23-Mar-98	0.01	0.04	0.56	0.52	0.20	0.53	0.89	0.74	0.34	0.90	0.19
24-Mar-98	0.00	0.04	0.44	0.36	0.20	0.42	0.79	0.74	0.43	1.00	0.18
25-Mar-98	0.01	0.04	0.36	0.24	0.27	0.16	0.89	0.74	0.34	0.94	0.13
26-Mar-98	0.01	0.04	0.38	0.32	0.27	0.26	0.83	0.84	0.33	0.92	0.09
27-Mar-98	0.01	0.04	0.38	0.32	0.27	0.32	0.76	0.84	0.33	0.89	0.08
28-Mar-98	0.01	0.04	0.36	0.44	0.20	0.21	0.77	0.84	0.34	0.75	0.27
29-Mar-98	0.01	0.04	0.26	0.32	0.29	0.11	0.61	0.74	0.48	0.90	0.30
Category 19											
30-Mar-99	5	0.5	147	118	48	57	42.8	0.22	179	8.08	13.0
01-Apr-99	6	0.5	150	126	47	56	41.6	0.23	171	8.08	15.2
02-Apr-99	8	0.5	161	130	49	60	40.6	0.22	170	7.96	14.8
Normalised values											
30-Mar-99	0.01	0.04	0.31	0.28	0.51	0.74	1.00	0.47	0.30	0.85	0.32
01-Apr-99	0.02	0.04	0.38	0.60	0.49	0.68	0.97	0.53	0.16	0.86	0.38
02-Apr-99	0.03	0.04	0.67	0.76	0.53	0.89	0.95	0.47	0.15	0.62	0.37

Category 20

DATE	Raw Turb Lab	TEMP (°C)	TOTAL HARD mg/L	TOTAL ALK. mg/L	Alum Plant #4	Lime Plant #4	Carbon Plant #4	Polymer Plant #4	Raw Flow	Recarb Effl pH	Res. Inf. particle counts (p/mL)
09-Apr-98	47	0.5	139	114	50	58	10.1	0.25	180	8.0	16.1
14-Apr-98	48	0.8	138	111	43	62	7.5	0.21	180	7.9	10.8
15-Apr-98	78	0.8	142	117	50	53	6.1	0.21	180	8.0	21.0
16-Apr-98	209	1	145	120	64	60	3.9	0.19	180	8.0	17.8
17-Apr-98	48	0.8	138	111	43	62	7.5	0.21	180	8.0	25.7
18-Apr-98	153	2.6	147	121	77	59	4.4	0.16	186	7.8	16.4
19-Apr-98	101	4.9	148	122	61	59	4.8	0.13	189	7.7	25.2
Normalised values											
09-Apr-98	0.22	0.04	0.10	0.12	0.54	0.79	0.24	0.63	0.31	0.70	0.41
14-Apr-98	0.22	0.07	0.08	0.00	0.42	1.00	0.18	0.42	0.31	0.58	0.28
15-Apr-98	0.37	0.07	0.18	0.24	0.54	0.53	0.14	0.42	0.31	0.69	0.54
16-Apr-98	1.00	0.10	0.26	0.36	0.78	0.89	0.09	0.32	0.31	0.67	0.45
17-Apr-98	0.22	0.07	0.08	0.00	0.42	1.00	0.18	0.42	0.31	0.61	0.67
18-Apr-98	0.73	0.27	0.31	0.40	1.00	0.84	0.10	0.16	0.41	0.22	0.42
19-Apr-98	0.48	0.51	0.33	0.44	0.73	0.84	0.11	0.00	0.46	0.00	0.66

Apr. 14 noise pattern lower than rest of category

Apr. 16 removed because calculated Z-score > 2.5 (due to high raw turbidity) as described in section 3.3.4, therefore pattern is an outlier

3.3.4 Data Set Extraction Preparation

Before models are to be trained, the data set of each sub-category was to be extracted into a training set, a testing set, and a production set. The training set is the set of data that the model is trained with, and it is tested on the testing set during training. The production set is data that is not presented to the model during training. Instead, it is used on the model after training is complete to determine the accuracy and robustness of the model. Variability of the data within each sub-category data set should no longer be too extensive as a result of categorisation. However, to ensure that distribution of patterns was even between each set, preparation of the data set was required. Essentially, the n-dimensional distance of each data pattern was calculated relative to the mean of the data set, and each pattern was subsequently ranked from furthest distance to the mean to the closest distance to the mean. Finally, data were extracted in that every third pattern was placed in the testing set and every fifth pattern was placed in the production set.

N-dimensional distance between two points is calculated using the following formula:

$$D_{AB} = \sqrt{(X_A - X_B)^2 + (Y_A - Y_B)^2 + \ldots + (Z_A - Z_B)^2} \qquad (7)$$

where, D_{AB} is the n-dimensional distance between pattern A and B,

X, Y,..., Z are values of each input for either pattern A or B.

Before the n-dimensional distance of a data pattern is calculated, the data set needs to be normalised, which scales the range of all input values between 0 and 1. The purpose of this is to ensure that each input value has an equal contribution or impact on determining the n-dimensional distance. Therefore, a significant change in one input will not be masked by changes in other inputs. Normalisation of an input value is calculated in the following way:

$$X_{norm(A)} = \frac{(X_A - X_{\min})}{(X_{\max} - X_{\min})} \quad (8)$$

where, $X_{norm(A)}$ is the normalized value of X_A

X_A is the actual value of input X of pattern A

X_{min} is the minimum value of input X in the data set

X_{max} is the maximum value of input X in the data set.

Once this was calculated for each input value in each sub-category data set, the n-dimensional distance relative to the mean of the data set was calculated for each pattern. The next step was to rank each pattern. In order to do that, the z-score of each pattern was calculated utilising the n-dimensional distance in the following way:

$$Zscore_A = \frac{(D_{norm(A)} - D_{norm(avg)})}{D_{norm(std.dev.)}} \quad (9)$$

where, $Zscore_A$ is the Z-score of pattern A

$D_{norm(A)}$ is the normalized n-dimensional distance of pattern A

$D_{norm(avg)}$ is the average normalized n-dimensional distance of the data set

$D_{norm(std.dev.)}$ is the standard deviation of the data set of normalized n-dimensional distances.

Once ranked and listed in descending order, every third pattern was placed in the testing set, every fifth pattern was placed in the production set, and the rest was placed in the training set.

3.4 Model Development and Optimisation

With the data sets now prepared, model development is the next step. Using the software NeuroShell 2 version 4.0, models were set up and trained. Training criteria were kept constant in all modelling. Pattern selection was set at random, weights were updated by momentum, and training was stopped either if the number of training epochs reached 10000 or if the average error for the test set reached 0.002. Model performance was measured by the R^2 value, the coefficient of multiple determination, and by the mean squared error, both of which were calculated by the software. The appendix section on pages 145-146 describes how these were calculated.

3.4.1 Architecture

The architecture of a neural network is the overall structure of the network, which includes the number of hidden layers and the types of connections between the layers. Backpropagation networks were the type of networks to be used, however there are four types of backpropagation networks to choose from with three variations of each for a total of twelve choices in the NeuroShell 2 software. Standard net is the first type in which each layer is connected only to the previous layer. Variations include three layers (one hidden layer), four layers (two hidden layers), and five layers (three hidden layers). The next type of backpropagation network is the Jordan-Elman net. These are recurrent networks with dampened feedback. Variations include input layer feedback, hidden layer feedback, and output layer feedback. The third type of backpropagation network is the Ward net, which has multiple hidden slabs with different activation functions. Variations include two hidden slabs with different activation functions, two hidden slabs with different activation functions and a jump connection, and three hidden slabs with different activation functions. The final type of backpropagation network is the jump connection net where each layer is connected to every previous layer. Variations here include three, four, or five layers. Each type of network was tried and tested with the results presented in Table 3-10.

Table 3-10. Model performance results testing various network architectures.

Type of Network	R-squared value				Mean Squared Error			
	Production	Test	Training	Entire	Production	Test	Training	Entire
5-layer standard net	**0.67**	**0.66**	**0.47**	**0.56**	**13.6**	**9.3**	**20.5**	**16.0**
4-layer standard net	0.64	0.62	0.45	0.53	14.8	10.3	21.2	16.9
3-layer standard net	0.63	0.58	0.41	0.50	15.6	11.3	23.0	18.2
recurrent net, input layer feedback	0.55	0.60	0.47	0.50	18.9	10.7	20.6	18.3
recurrent net, hidden layer feedback	0.54	0.56	0.47	0.51	19.3	11.8	20.6	17.9
recurrent net, output layer feedback	0.58	0.57	0.46	0.51	17.4	11.7	21.0	17.6
ward net, 2 hidden slabs, no jump connection	**0.72**	**0.57**	**0.57**	**0.60**	**11.6**	**11.6**	**16.9**	**14.4**
ward net, 3 hidden slabs	0.64	0.60	0.40	0.50	15.1	10.8	23.4	18.2
ward net, 2 hidden slabs, with jump connection	0.56	0.59	0.50	0.54	18.1	11.0	19.2	16.6
3-layer jump connection net	0.56	0.60	0.46	0.52	18.1	10.7	21.1	17.5
4-layer jump connection net	0.48	0.52	0.50	0.51	21.6	13.0	19.6	18.0
5-layer jump connection net	0.51	0.56	0.48	0.51	20.2	11.9	20.2	17.8

Default settings were used in running these models, and the number of hidden neurons were set to be twice the number of input neurons in each hidden layer. The results from the table are from the category 1, sub-cat 1 model. It can be seen that the two best results are the 5-layer standard net and the ward net with two hidden slabs (no jump connection). Since the standard net is a common and virtually universal architecture for most problem domains, it was decided to use the 5-layer standard net exclusively.

3.4.2 Modelling Parameters

After determining the best architecture, other parameters can be examined and optimised. These parameters include the scale function, activation function, number of hidden neurons, learning rate, momentum, and initial weights. The category 1, sub-cat 1 model was used in examining each parameter.

One of the first parameters tested was the number of hidden neurons. With three layers of hidden neurons, many combinations of numbers are available, thus multiple arbitrary combinations were tested. Table 3-11 shows what was tested and the results.

Table 3-11. Model performance results testing various combinations of numbers of hidden neurons.

5-layer bp neuron configuration	R-squared value				Mean Squared Error			
	Production	Test	Training	Entire	Production	Test	Training	Entire
10-10-10-10-1	0.66	0.65	0.46	0.55	14.2	9.5	21.0	16.4
10-20-20-20-1	**0.67**	**0.66**	**0.47**	**0.56**	**13.6**	**9.3**	**20.5**	**16.0**
10-30-30-30-1	0.65	0.62	0.45	0.53	14.7	10.3	21.3	16.9
10-20-10-10-1	0.67	0.63	0.45	0.54	13.5	10.0	21.2	16.6
10-10-20-10-1	0.65	0.64	0.46	0.55	14.4	9.6	20.8	16.4
10-10-10-20-1	0.64	0.64	0.45	0.54	15.0	9.7	21.2	16.7
10-20-20-10-1	0.65	0.62	0.46	0.54	14.5	10.2	20.9	16.6
10-10-20-20-1	0.70	0.60	0.40	0.51	12.5	10.8	23.3	17.8
10-15-15-15-1	0.67	0.65	0.45	0.55	13.7	9.5	21.2	16.5
10-25-25-25-1	0.61	0.59	0.45	0.52	16.3	11.0	21.2	17.4

Default settings were used in each model. Out of the combinations that were tested, the configuration of twenty neurons in each hidden layer had the best overall results when tested using the production set, training set, testing set and the entire data set.

The next parameter tested and examined was the activation function of the hidden layers with eight different activation functions available to choose from in NeuroShell 2. These include logistic, symmetric logistic, linear, tanh, tanh15, Gaussian, Gaussian complement, and sine. With a 5-layer standard net being used, there are three hidden layers whose activation function can be changed. Default settings for initial weights, momentum, learning rate and scaling function was used. After some extensive testing with various combinations, it was found that the best statistical results were obtained with a logistic function set for the first hidden layer, a Gaussian function for the second hidden layer, and a tanh15 function for the third hidden layer compared to the default setting of logistic function for each of the three hidden layers. This is shown in Table 3-12. However, problems arose with the actual predictions of the model.

Table 3-12. Best model performance results from the testing of various combinations of activation functions.

Activation function	R-squared value				Mean Squared Error			
	Production	Test	Training	Entire	Production	Test	Training	Entire
logistic	0.68	0.61	0.45	0.54	13.8	9.5	18.4	14.9
Gaussian / tanh15	0.76	0.63	0.47	0.59	9.9	9.3	18.3	14.2

As can be seen in Figure 3-8 below, which is production set data, network predictions are fixated around two points: 18 particle counts/mL and 9 particle counts/mL. This is even more apparent when running the entire data through the model as shown in Figure 3-9. In contrast, this behaviour was not observed in the model using the logistic function solely as the activation function, and this is demonstrated in Figure 3-10 using production set data and Figure 3-11 using the entire data set.

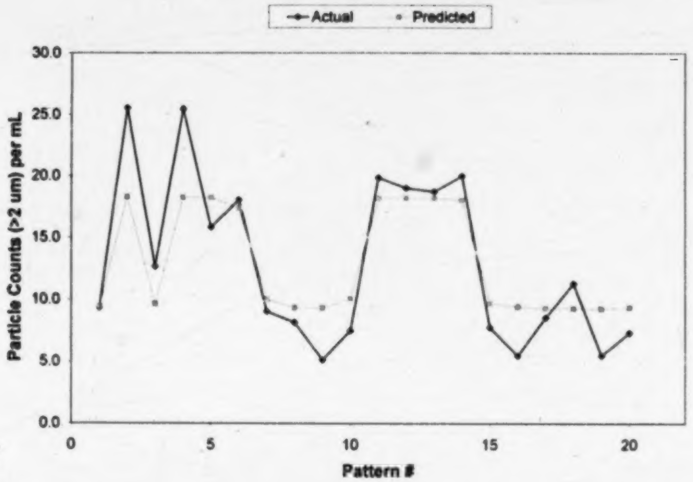


Figure 3-8. Actual vs. model predicted particle counts using a Gaussian and tanh 15 activation function and production set data.

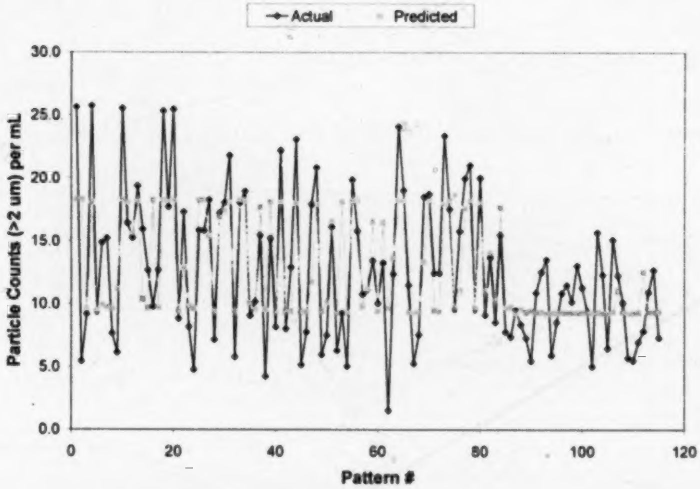


Figure 3-9. Actual vs. model predicted particle counts using a Gaussian and tanh 15 activation function and the entire cat 1, sub-cat 1data set.

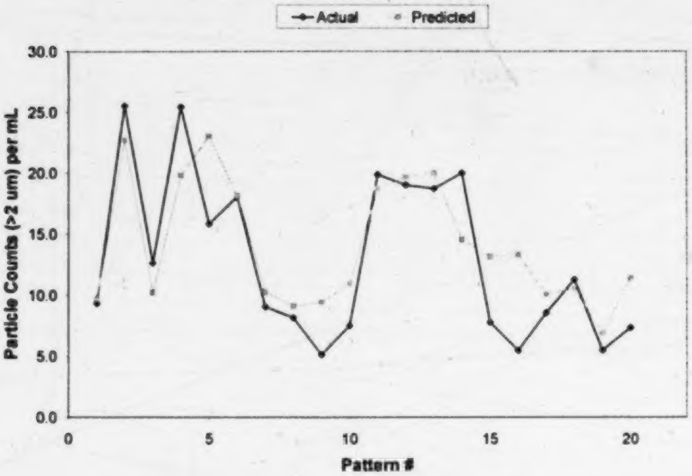


Figure 3-10. Actual vs. model predicted particle counts using the logistic activation function and production set data.

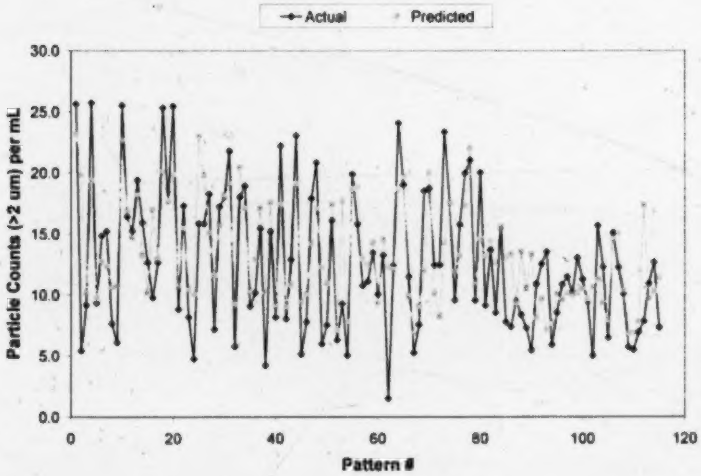


Figure 3-11. Actual vs. model predicted particle counts using the logistic activation function and the entire cat 1, sub-cat 1data set.

The main reason for these results is the different mapping characteristics of each activation function. The Gaussian function maps values between 0 and 1, therefore the output of that function will always be positive. The Tanh15 function maps positive values between 0 and 1 while negative values are mapped between –1 and 0. Since these two activation functions are quite different, using them in subsequent order is what is causing the network predictions to fixate around the two values illustrated in Figures 3-8 and 3-9. Using one consistent activation function such as the logistic function does not result in this problem as illustrated in Figures 3-10 and 3-11. Therefore, it was decided to keep using the logistic function, which is a commonly used in many neural network applications, as the activation function in all models in each hidden layer.

Learning rate, momentum, and initial weights of the links between each layer are other factors that can be modified. Each of the four links was changed individually at first. Initial weights were set at 0.3 or 0.7. Learning rate and momentum were set at combinations of 0.1 or 0.7 as well as setting both at only 0.4. Two, three, and four links were then changed at a time based on the best results of changing each link by itself. Table 3-13 shows the statistical results of the optimal combination compared to the default setting. As it turned out, the best overall result was obtained by setting link 3 with a learning rate and momentum of 0.4 and the initial weight at 0.7 while keeping the other links at default settings. All the results are shown in the appendix on pages 146-149.

Table 3-13. Best model performance result after testing various combinations of learning rate, momentum, and initial weights.

	R-squared value				Mean Squared Error			
	Production	Test	Training	Entire	Production	Test	Training	Entire
all links (default) learning rate = 0.1, momentum = 0.1, initial wt = 0.3	0.68	0.61	0.45	0.54	13.8	9.5	18.4	14.9
link 3 learning rate = 0.4, momentum = 0.4, initial wt = 0.7 links 1, 2, 4 = default	0.77	0.64	0.47	0.58	9.8	9.2	18.1	14.1

The final parameter to investigate is the scaling function for which there are seven options available in NeuroShell 2. These include linear functions scaled from zero to one or minus one to one. Each of these ranges can be set with closed brackets, [], or open brackets, <<>>. The closed brackets mean that later data values beyond a given range are scaled to 0 or 1 or -1, whereas the open brackets would allow data beyond a given range to be scaled beyond 0, 1, or –1. The remaining options for scaling functions include a logistic function, a tanh function, both of which are non-linear, or no scaling function at all. Each scaling function was tested and it was found that there was no significant difference in model performance between each function as shown in Table 3-14. Using no scaling function was the only option that resulted in poor performance. Therefore, it was decided to use a linear function that scaled <<-1,1>>, which provides a larger range than 0,1, and the open brackets allow for future data that may go beyond the defined range to be scaled more accurately and relevantly.

Table 3-14. Model performance results of testing various scale functions.

scale function for slab 1	R-squared value				Mean Squared Error			
	Production	Test	Training	Entire	Production	Test	Training	Entire
linear [-1,1]	0.77	0.64	0.47	0.58	9.8	9.2	18.1	14.1
linear <-1,1>	0.77	0.64	0.47	0.58	9.8	9.2	18.1	14.1
linear [0,1]	0.80	0.63	0.45	0.57	8.5	9.6	18.7	14.3
linear <0,1>	0.80	0.63	0.45	0.57	8.5	9.6	18.7	14.3
logistic	0.81	0.60	0.49	0.59	8.2	10.3	17.4	13.7
tanh	0.78	0.55	0.47	0.56	9.06	11.5	18.2	14.7
none	-0.07	0.02	-0.03	-0.02	44.6	25.1	35.2	33.8

3.4.3 Model Amalgamation

With parameters set at optimal settings, the next step is to combine the resulting models from each sub-category. The sub-category models within category 1 were to be combined as well as the sub-category models within category 2. Therefore, the data set for each model to be combined was made into one data set. Then, extraction into the training, testing, and production set was done with every third pattern placed in the testing set and every fifth pattern placed in the production set with the rest going into the training set. A model was then developed using this combined data set and using the same settings as before for the various modelling parameters. Tables 3-15 and 3-16 show model performance results of the models for category 1 and 2 using the different data sets to test performance compared to the results of combining the sub-category models. As shown in Table 3-15, combining the category 1 sub-category models resulted in a model that performed more poorly than the separate sub-category models. Therefore, it was decided to keep them separate as the category 1, sub-cat 1 (C1S1) model and category 1, sub-cat 2 (C1S2) model. On the other hand, Table 3-16 shows that the combined model outperforms the category 2, sub-cat 1 (C2S1) model and is comparable to the category 2, sub-cat 2 (C2S2) model. Therefore, it was decided to keep the two sub-category models as one combined model, the category 2 model (C2).

Table 3-15. Model performance results of the category 1 sub-category models compared to the combined model.

		C1S1 model	C1S2 model	combined model
R squared	production set	0.80	0.51	0.33
	test set	0.59	0.18	0.09
	training set	0.40	0.28	0.26
	entire set	0.51	0.29	0.22
mean squared error	production set	8.7	11.0	23.4
	test set	10.5	24.5	42.2
	training set	22.9	37.2	27.2
	entire set	17.5	29.0	30.5

Table 3-16. Model performance results of the category 2 sub-category models compared to the combined model.

		C2S1 model	C2S2 model	combined model
R squared	production set	0.29	0.54	0.42
	test set	0.11	0.05	0.24
	training set	0.34	0.70	0.78
	entire set	0.22	0.38	0.49
mean squared error	production set	7.2	7.0	5.8
	test set	41.5	36.7	24.5
	training set	7.3	3.6	3.4
	entire set	16.6	12.7	9.8

At this point, data that were set aside during categorisation and noise analysis were inserted into the relevant data set and the model was retrained. If the insertion of these additional patterns did not affect model performance too adversely, they were kept in the model. R^2 values and mean squared error were evaluated and compared using not only the production set, but also the training, testing and entire set for evaluation just as for evaluating the combined models as listed in Tables 3-15 and 3-16. If there was significant overall deterioration in model performance, the patterns were removed. Tables 3-17, 3-18, and 3-19 list performance results of the C1S1, C1S2, and C2 models when noise patterns and extreme case patterns were included in the model data set. As it can be seen, performance generally deteriorated in most cases. However, the case in which four extreme patterns were included in the C1S1 model data set did not show significantly negative impacts on model performance. In fact, convergence improved slightly as shown by the slightly higher R^2 and lower mean squared error values when using the training and testing set to test model performance. Therefore, these four data patterns were kept as part of the C1S1 model data set.

Table 3-17. C1S1 model performance results with the inclusion of noise patterns and extreme case patterns.

		C1S1 model	6 noise patterns added	11 noise patterns added	**4 extreme patterns added**	10 extreme patterns added	17 extreme patterns added
R squared	production set	0.80	0.50	0.53	**0.79**	0.60	0.62
	test set	0.59	0.43	0.45	**0.60**	0.61	0.63
	training set	0.40	0.39	0.34	**0.45**	0.40	0.45
	entire set	0.51	0.42	0.39	**0.54**	0.48	0.53
mean squared error	production set	8.7	22.5	20.6	**8.9**	17.0	16.6
	test set	10.5	15.8	15.0	**10.1**	10.6	11.6
	training set	22.9	27.1	35.5	**21.3**	22.6	20.3
	entire set	17.5	23.5	28.0	**16.6**	18.6	17.4

Table 3-18. C1S2 model performance results with the inclusion of noise patterns and extreme case patterns.

		C1S2 model	6 noise patterns added	12 noise patterns added	6 extreme patterns added
R squared	production set	0.51	0.30	-0.39	0.13
	test set	0.18	0.35	0.35	0.07
	training set	0.28	0.52	0.64	0.11
	entire set	0.29	0.38	0.44	0.10
mean squared error	production set	11.0	23.3	106.1	28.5
	test set	24.5	41.1	258.0	57.7
	training set	37.2	12.4	141.1	23.3
	entire set	29.0	25.2	167.7	35.9

Table 3-19. C2 model performance results with the inclusion of noise patterns and extreme case patterns.

		C2 model	8 noise patterns added	16 noise patterns added	2 extreme patterns added	10 extreme patterns added
R squared	production set	0.42	0.17	-0.04	-0.10	0.13
	test set	0.24	0.02	0.00	0.00	0.28
	training set	0.78	0.22	0.00	-0.03	0.71
	entire set	0.49	0.15	0.00	-0.02	0.46
mean squared error	production set	5.8	11.3	21.4	10.9	9.0
	test set	24.5	34.2	39.1	32.4	21.9
	training set	3.4	20.2	42.8	16.0	4.5
	entire set	9.8	22.6	38.1	19.7	10.3

4.0 Results

At the end of model development and optimisation, three models were produced; models for sub-category 1 and 2 of category 1 (C1S1 and C1S2) and a model for category 2 (C2). Section 4.1 presents performance results of these models while examples of predictions of each model as a result of varying different input values are presented in section 4.2. In addition, the significance of the prediction results will also be discussed. Performance was measured by the R^2 value, mean squared error, and mean and maximum absolute error. The appendix on pages 145-146 describes each parameter and how it was calculated by the software.

4.1 Model Performance Results

Table 4-1 shows the statistics of the data set used for the C1S1 model, which contains 115 patterns in total. The vast majority of patterns come from the months of October, November, March, and April corresponding with fall and spring. This is also reflected in the observations of the average raw turbidity and temperature being slightly higher than the average for the data set of the C1S2 model shown in Table 4.3. PAC dose is also significantly higher. Therefore, the C1S1 model will be referred to as the spring/fall model.

The spring/fall model is a 5-layer backpropagation network with twenty hidden neurons in each hidden layer, ten input neurons, and one output neuron. A linear scale function that scales <<-1,1>> as explained in section 3.4.2 was used along with a logistic activation function. Each link between each layer was set with a learning rate of 0.1, a momentum of 0.1, and an initial weight of 0.3, except for link 3, which links slab 3 and 4, where learning rate is set to 0.4, momentum to 0.4 and initial weight to 0.7.

Table 4-1. Spring/fall (C1S1) data set characteristics.

	Mean	Range	Std. Dev.
Raw Turbidity (NTU)	14.6	2 - 172	29.3
Temperature (°C)	2.3	0.1 - 9.5	2.6
Total Hardness (mg/L)	156.4	134 - 176	8.7
Total Alkalinity (mg/L)	124.8	110 - 138	6.1
Alum Dose (mg/L)	36.6	18 - 117	15.4
Lime Dose (mg/L)	53.8	43 - 81	7.4
PAC dose (mg/L)	7.9	0 - 44.6	12.5
Polymer Dose (mg/L)	0.23	0.12 - 0.36	0.04
Raw Flow (ML/d)	190.4	149 - 222	16.3
pH	7.93	7.64 - 8.15	0.13
Reservoir Influent Particle Counts (counts/mL)	13.8	1.5 - 27.9	6.0

Table 4-2. Spring/fall (C1S1) model performance results.

R^2	0.79
Mean Squared Error	8.9
Mean Absolute Error	2.3 counts/mL
Maximum Absolute Error	6.7 counts/mL

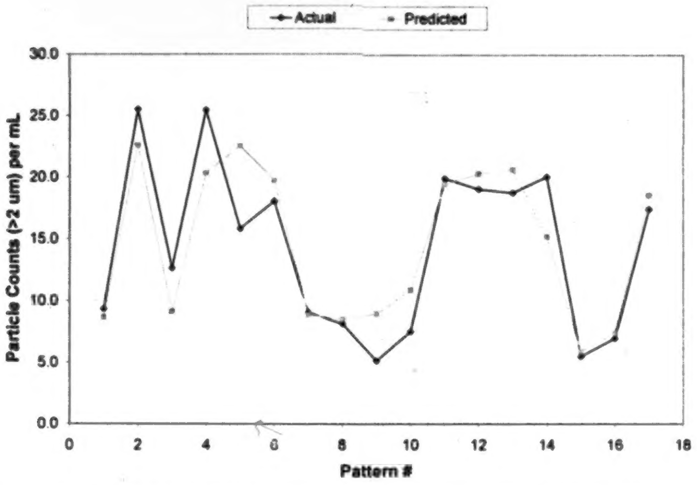


Figure 4-1. Spring/fall (C1S1) model predictions vs. actual particle count data.

Table 4-2 shows the statistical results of the spring/fall model performance based on the production set. The R^2 value is quite good and errors are quite low. Figure 4-1 compares model predictions to actual particle count data, which are from the production set. As can be seen, trends in the data tend to be followed well by the model.

88 data patterns were used as the data set for the C1S2 model. Table 4-3 shows the statistics of this data set. The majority of patterns come from the months of December, January, and February, which corresponds with winter, and this is associated with consistently low raw turbidity and low temperatures. Therefore, the C1S2 model will be referred to as the winter model from this point onward.

Table 4-3. Winter (C1S2) data set characteristics.

	Mean	Range	Std. Dev.
Raw Turbidity (NTU)	2.6	2.0 - 5.0	0.8
Temperature (°C)	0.53	0.5 - 0.6	0.05
Total Hardness (mg/L)	171.0	156 - 184	7.3
Total Alkalinity (mg/L)	131.3	119 - 144	5.0
Alum Dose (mg/L)	34.3	31 - 39	1.2
Lime Dose (mg/L)	62.6	52 - 71	5.1
PAC dose (mg/L)	0.04	0 - 2.40	0.28
Polymer Dose (mg/L)	0.33	0.26 - 0.43	0.04
Raw Flow (ML/d)	170.4	156 - 199	10.0
pH	7.89	7.62 - 8.05	0.11
Reservoir Influent Particle Counts (counts/mL)	11.8	2.2 - 40.0	6.4

The winter model is a 5-layer backpropagation network with twenty hidden neurons in each of the three hidden layers, ten input neurons, and one output neuron. A logistic activation function was used for each hidden layer, and a linear scale function in the input layer that scales <<-1,1>> was also used. All links between each layer were set with a learning rate of 0.1, a momentum of 0.1, and an initial weight of 0.3.

Table 4-4. Winter (C1S2) model performance results.

R^2	0.51
Mean Squared Error	11.0
Mean Absolute Error	2.8 counts/mL
Maximum Absolute Error	6.9 counts/mL

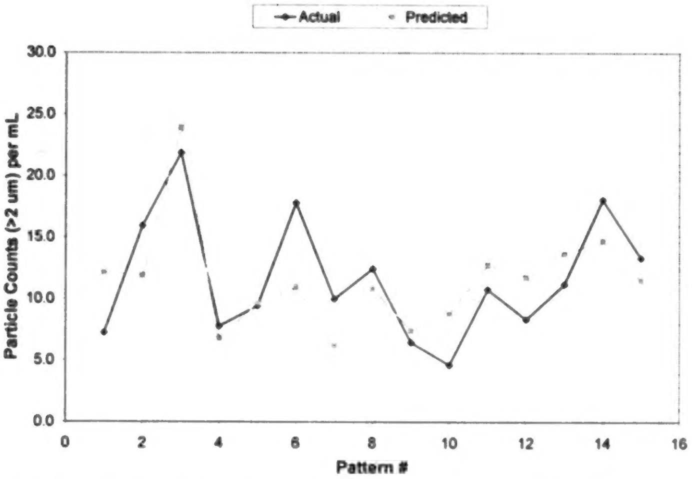


Figure 4-2. Winter (C1S2) model predictions vs. actual particle count data.

Table 4-4 shows the statistical results of the winter model performance based on the production data set. The R^2 value is somewhat low, but the errors are quite low. Figure 4-2 compares model predictions to actual particle count data from the production set. Although peaks in the data were missed in some cases, the general trends in the data still tend to be followed by the model.

Statistical characteristics of the data set of 137 patterns used for the C2 model is shown in Table 4-5. Data patterns generally come from between May and September, which corresponds with summer weather. The main difference in the data set compared to category 1 is that average temperature and raw turbidity are higher in category 2 than in category 1. In addition, the range of raw turbidity encountered in category 2 is higher than the range of the data sets in category 1. Therefore, the C2 model will be referred to as the summer model from this point onward.

Table 4-5. Summer (C2) data set characteristics.

	Mean	Range	Std. Dev.
Raw Turbidity (NTU)	51.1	2 - 720	98.7
Temperature (°C)	17.1	9.1 - 24.6	3.6
Total Hardness (mg/L)	154.3	140 - 181	7.8
Total Alkalinity (mg/L)	126.3	113 - 143	5.8
Alum Dose (mg/L)	45.0	23 - 124	22.5
Lime Dose (mg/L)	54.4	38 - 90	11.0
Polymer Dose (mg/L)	0.17	0.12 - 0.22	0.02
Raw Flow (ML/d)	200.5	100 - 253	22.6
pH	8.02	7.66 - 8.35	0.18
Reservoir Influent Particle Counts (counts/mL)	7.8	2.5 - 26.3	4.4

The summer model is a 5-layer backpropagation network with eighteen hidden neurons in each of the three hidden layers and a logistic function as the activation function for each hidden layer. The scale function in the input layer is a linear function that scales <<-1,1>>, and links between each layer was set with a learning rate of 0.1, a momentum of 0.1 and an initial weight of 0.3. The summer model also differs from the spring/fall and winter models in that there are nine inputs in the summer model compared to ten in the other two models. This is because PAC dose was not included as an input. Only two patterns out of the 137 in total contained a value of PAC other than 0 mg/L, therefore it was decided not to include this parameter as an input. The output parameter of reservoir influent particle counts remains the same, but eighteen hidden neurons were used instead of twenty to follow an arbitrary rule of two times the number of input neurons for each hidden layer used in the spring/fall and winter models.

Table 4-6. Summer (C2) model performance results.

R^2	0.42
Mean Squared Error	5.8
Mean Absolute Error	2.0 counts/mL
Maximum Absolute Error	5.4 counts/mL

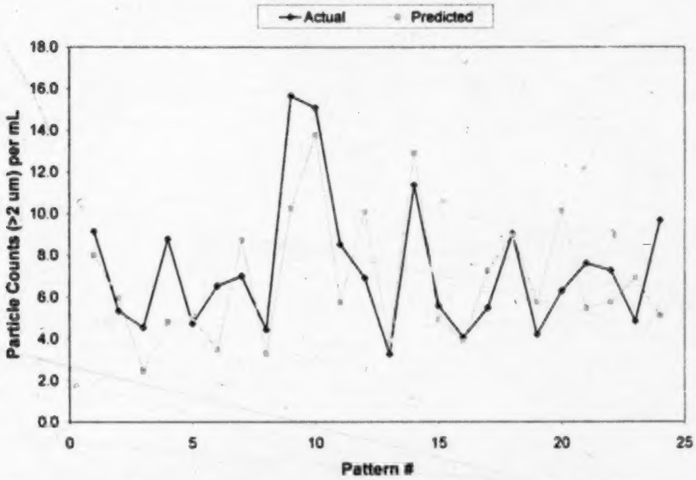


Figure 4-3. Summer (C2) model predictions vs. actual particle count data.

Table 4-6 shows the statistical results of the summer model performance based on production set data. Although the R^2 value is low, the absolute errors are low as well. Figure 4-3 shows how well the model predicts compared to the actual particle count values. Some peaks are missed and sometimes trends are not followed precisely, which accounts for the relatively low R^2 value.

4.2 Model Predictions

Now that the models have been trained and developed, they can be used to determine the impact any of the parameters used in the model has on the reservoir influent (combined filter effluent) particle counts. Presented in this section are the results of model predictions made by each model in response to varying different inputs. From these results, the effect of the varied input on particle counts can be examined. The difference of this effect under different seasonal conditions can be examined as well since each model represents different seasonal conditions with the C1S1 model representing spring and fall, the C1S2 model representing winter, and the C2 model representing summer. Section 4.2.1 presents results of varying a single input and comparisons are made between the three models. Section 4.2.2 presents the resulting effects of varying two parameters at a time.

In order to use the models to make predictions, a dynamic link library was created for each model, which allows for the model to be called and utilised from a spreadsheet program such as Microsoft Excel. Therefore, using the spreadsheet as an interface, values of each input can be typed onto the spreadsheet, and then these values are fed through a given model with the model predicted output written to the spreadsheet.

Listed in Table 4-7 are the values of each input when that input was held constant while varying one or two other inputs. For example, if model predictions were made with the spring/fall model while varying alum dose, raw turbidity would be set at 10 NTU, temperature at 2°C, hardness at 160 mg/L, alkalinity at 125 mg/L, lime dose at 55 mg/L, PAC dose at 0 mg/L, polymer dose at 0.25 mg/L, raw flow at 190 ML/d, and pH at 7.90. These set values were chosen based on the statistical average of each model's data set (shown in Tables 4-1, 4-3, and 4-5) with the values chosen as the average or near the average.

Table 4-7. Values of inputs held constant during examination of effects of inputs on particle counts.

	Spring/Fall (C1S1) model	Winter (C1S2) model	Summer (C2) model
Raw Turbidity (NTU)	10	2	50
Temperature (°C)	2	0.5	20
Total Hardness (mg/L)	160	170	155
Total Alkalinity (mg/L)	125	130	130
Alum Dose (mg/L)	35	35	50
Lime Dose (mg/L)	55	60	60
PAC dose (mg/L)	0	0	-
Polymer Dose (mg/L)	0.25	0.30	0.18
Raw Flow (ML/d)	190	170	200
pH	7.90	7.90	8.00

It should be noted, however, that model predictions are limited by the boundaries of the range of data values used to train the models. If presented with data beyond this range, model predictions are extrapolated and are not necessarily as accurate as opposed to predictions based on data within the trained range.

4.2.1 Model Predictions With One Varied Input

The parameters that were varied and examined for their effect on particle counts are alum dose, lime dose, PAC dose, and polymer dose. These factors were chosen based on the fact that these are parameters that can be modified and changed during operation of a WTP.

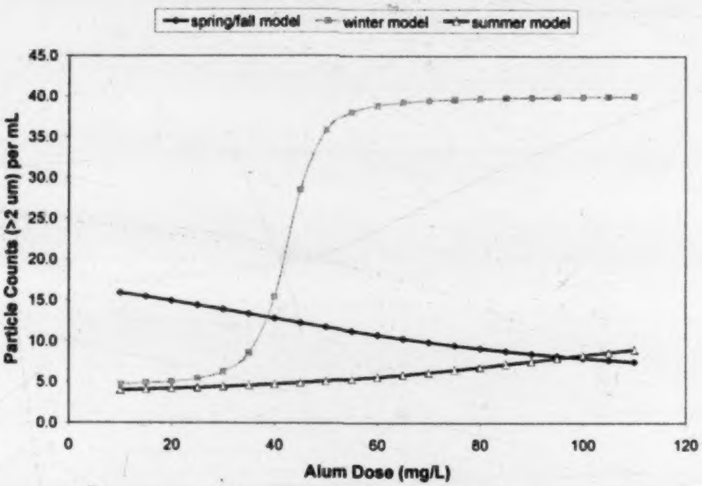


Figure 4-4. The effect of alum dose on particle counts as predicted by each model.

Shown in Figure 4-4 is the effect of alum dosage on particle counts predicted by each model. As it can be seen, the spring/fall model predicts that there is a decrease in particle counts as alum dosage is increased, however, the effect is not very big. The winter model predicts the opposite where increasing alum dose increases particle counts particularly when alum dose is greater than 35 mg/L. The reason for this is that under winter conditions, temperature and turbidity are quite low. With such a low turbidity, adding excess alum may be adding particles to the system. Moreover, lower temperatures decrease the efficiency of filtration as discussed in section 2.4.2.9. Therefore, these two factors combined may account for the observed increase in particle counts with increasing alum dosage. Another reason could be that excess alum is causing charge reversal in which particles result in having a positive charge from a negative charge rather than being neutralised. Consequently, coagulation/flocculation is less effective ultimately resulting in more particles passing through the filters. The summer model also predicts that as alum dose increases, particle counts increase but not by as much. This is likely due to the effect of adding particles to the system with excess loading of alum or possibly due to charge reversal as was described.

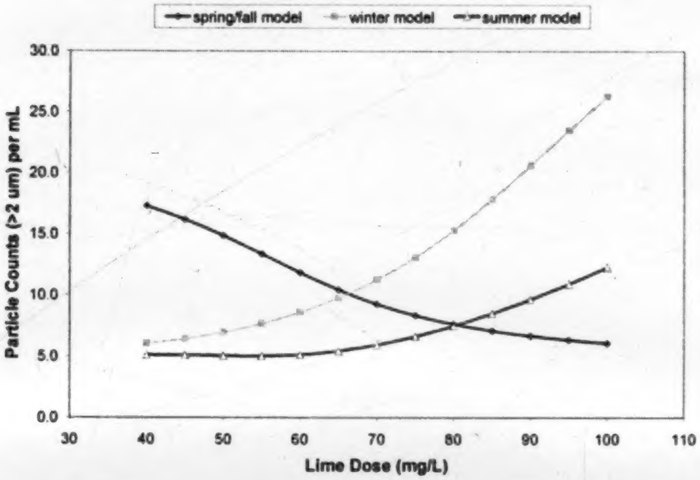


Figure 4-5. The effect of lime dose on particle counts as predicted by each model.

Figure 4-5 shows the effect of varying lime dose on particle counts while other inputs are held constant as predicted by each model. The trend predicted by the spring/fall model of decreasing particle counts with increasing lime dose could be attributed to extensive removal of calcium and magnesium particles associated with hardness. It is also possible that with high amounts of calcium carbonate precipitating, more particles that did not settle during coagulation/flocculation are being removed in secondary sedimentation by adhering to the calcium carbonate precipitate. With the winter model, it can been seen that there is a gradual increase in particle counts with an increase in lime dosage. The reason for this is likely similar to the one described for observations for alum dose. Although water is harder on average in the winter model than in the spring/fall model, the temperature is lower, and reaction kinetics are slower. Thus, calcium carbonate may not be precipitating out as fast as when temperatures are warmer resulting in particle formation beyond sedimentation and ultimately resulting in higher particle counts. For the summer model, it can be seen that generally, particle counts increase as lime dose increases at doses greater than 70 mg/L. The reason for this is likely due to particles being added to the system from excess dosages of lime.

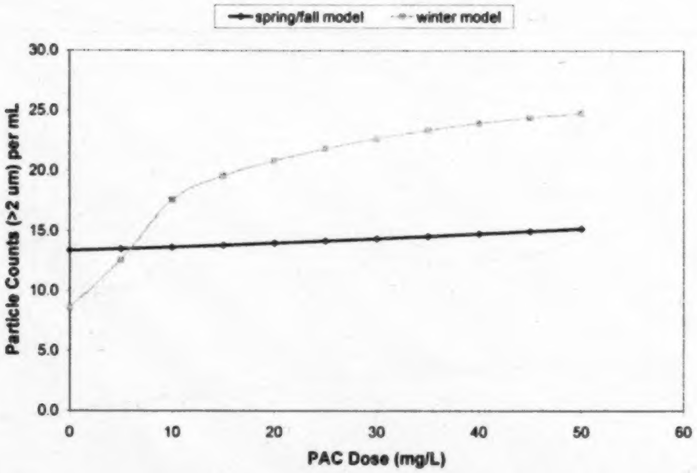


Figure 4-6. The effect of PAC dose on particle counts as predicted by each model.

Figure 4-6 shows the effect of PAC dose on particle counts while keeping the other input values constant. The summer model does not include PAC dose as an input as discussed in section 4.1, therefore it is not shown in the graph. It can be seen that increasing the PAC dose results in increased particle counts, minimally for the spring/fall model and more significantly for the winter model. Since particles are being added to the system, the extra particles detected are likely PAC particles, and adding more PAC therefore increases particle counts.

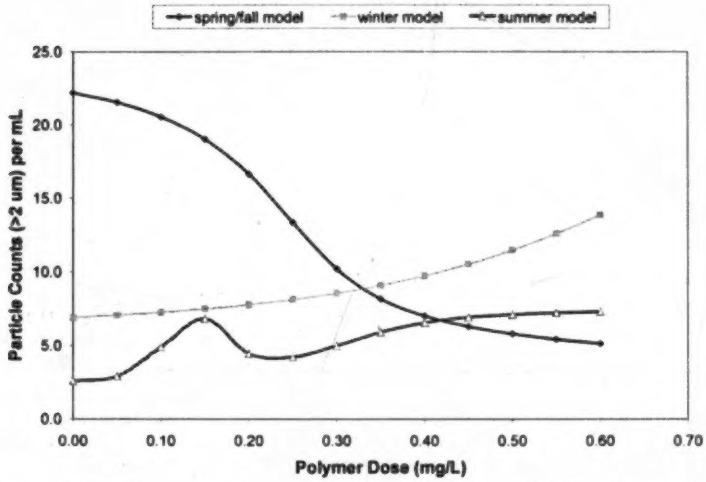


Figure 4-7. The effect of polymer dose on particle counts as predicted by each model.

The effect of polymer dose on particle counts as predicted by each model is shown in Figure 4-7. For the spring/fall model, it can be clearly seen that the use of polymers reduces particle counts in finished water. However, the curve levels off indicating that it is unnecessary to have high levels of polymer to further reduce particle counts below 5 counts/mL. The opposite is observed for the winter model where increased polymer dosage results in increased particle counts. The likely reason for this observation is similar to the one given for the trend observed with varying alum doses. With low turbidity, less coagulant is necessary and excess polymer added may be adding particles to the system. Together with lower filtration efficiency due to lower temperatures, the result would be an increase in particle counts with increased polymer dosage. The summer model shows a much different curve in which particle count levels peak at a dosage of 0.15 mg/L, decreases for a bit before increasing again and levelling off. The reason for this behaviour is unknown.

Table 4-8. Overview of the impact of parameters on finished water particle counts.

		Predicted Particle Count Range (counts/mL)		
Parameter	Parameter Range	Spring/Fall (C1S1) model	Winter (C1S2) model	Summer (C2) model
Alum Dose	10-110 mg/L	7-16	5-42	4-8
Lime Dose	40-100 mg/L	5-17	5-27	5-12
PAC Dose	0-50 mg/L	13-14	7-25	-
Polymer Dose	0-0.6 mg/L	5-22	7-15	3-7

Table 4-8 provides an overview of the extent of impact each of the four parameters that were examined has on finished water particle counts according to the models developed. Essentially, information from Figures 4-4 to 4-7 is summarised in Table 4-8. The parameter range indicates the range in which the parameters were varied. The predicted particle count range lists the range of particle counts that were predicted by each model within the parameter range.

Under spring and fall conditions (C1S1 model), polymer dose appears to have the largest impact on particle counts followed by alum and lime dose. PAC dose does not appear to have a relatively significant impact. Under winter conditions (C1S2 model), alum, lime, and PAC dose appear to have the most impact on particle counts with polymer dose having a relatively smaller impact. Under summer conditions (C2 model), alum, lime, and polymer dose does have an impact on particle counts, but not to the same extent as under spring, fall, or winter conditions.

4.2.2 Model Predictions With Two Varied Inputs

The effects of varying alum and lime dose plus alum and polymer dose at the same time on particle counts are presented in this section. These factors were chosen to be examined because they can be changed operationally in a WTP, therefore these results serve as a demonstration of the usefulness to operators of utilising these models in determining operating conditions. The other inputs are held constant as listed in Table 4-7.

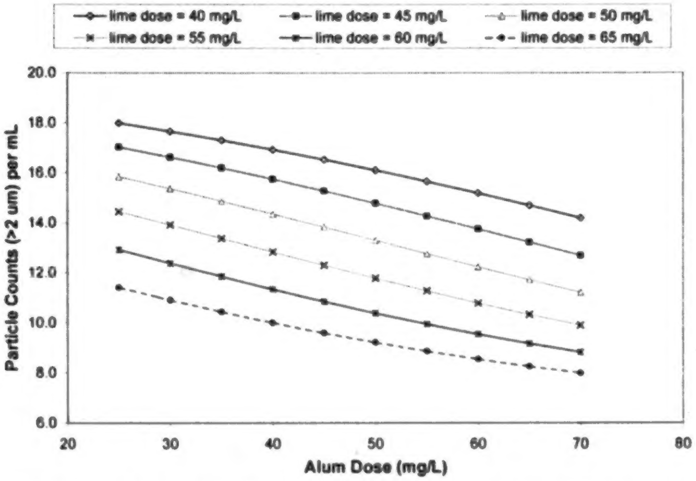


Figure 4-8. The effect of alum and lime dose on particle counts as predicted by the spring/fall model.

In Figure 4-8, the spring/fall model predicts the effect of alum dose and lime dose on particle counts while the other inputs are held constant. It is observed that particle counts decrease as alum dose and lime dose increases, and the trends are similar to those observed in Figures 4-4 and 4-5.

The same inputs are examined in Figure 4-9, but the winter model was used. It can be seen that with both increasing alum and lime dosage, particle counts increase due to the same reasons described in section 4.2.1.

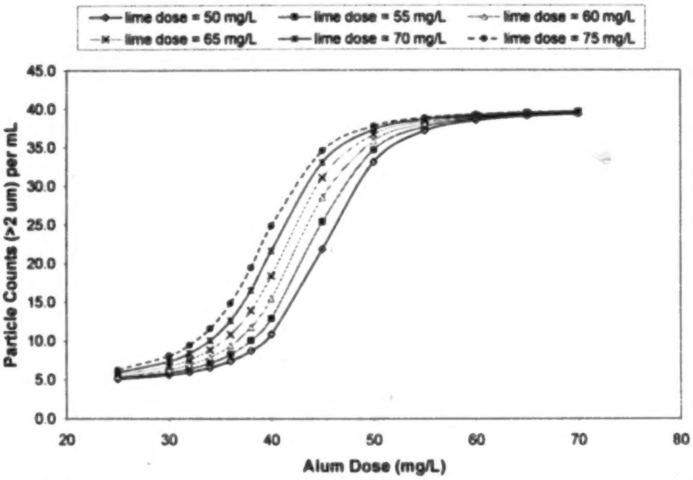


Figure 4-9. The effect of alum and lime dose on particle counts as predicted by the winter model.

Figure 4-10 also shows the effect of alum and lime dose but as predicted by the summer model.

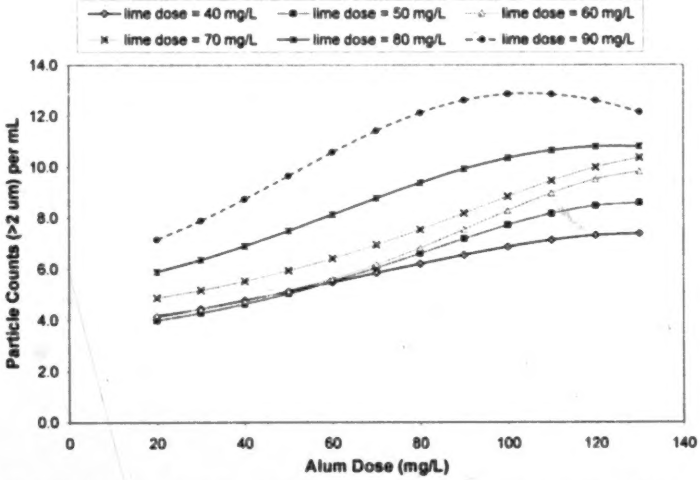


Figure 4-10. The effect of alum and lime dose on particle counts as predicted by the summer model.

It is generally seen in Figure 4-10 that increasing both alum and lime doses increases particle counts likely due to the effect of the addition of particles to the system at high doses of these chemicals.

An operator can use the type of information that can be generated such as those presented in Figures 4-8 to 4-10 to help determine the amount of chemicals to use. If the main goal is to reduce particle count levels, the operator can try different levels and combinations of alum and lime dose and determine from the model predictions the optimal combination of dosages. For example, under spring conditions, if an alum dose of 40 mg/L is to be used, a lime dose of 40 mg/L will result in 18 particle counts/mL or a lime dose of 65 mg/L would result in 10 particle counts/mL based on the spring/fall model in Figure 4-8. The operator could then decide which option is more optimal when factoring in chemical costs. However, under winter conditions shown in Figure 4-9, although lime dose has a minor effect on particle counts, the choice for alum dose needs to be considered carefully as too high a dose would result in high particle count levels. Summer conditions shown in Figure 4-10 indicate that different doses do not significantly vary the resulting water quality, therefore it is up to the operator to decide the economically optimal choice.

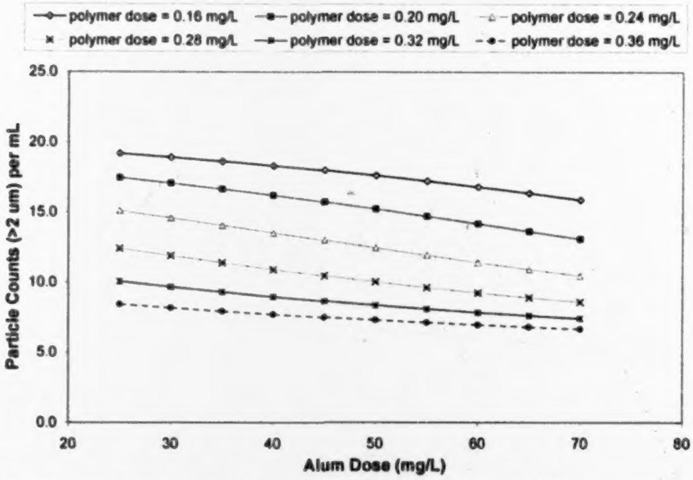


Figure 4-11. The effect of alum and polymer dose on particle counts as predicted by the spring/fall model.

Figure 4-11 shows the effect of alum and polymer dose on particle counts as predicted by the spring/fall model. The trend is apparent with a decrease in particle counts as alum dose and polymer dose increases.

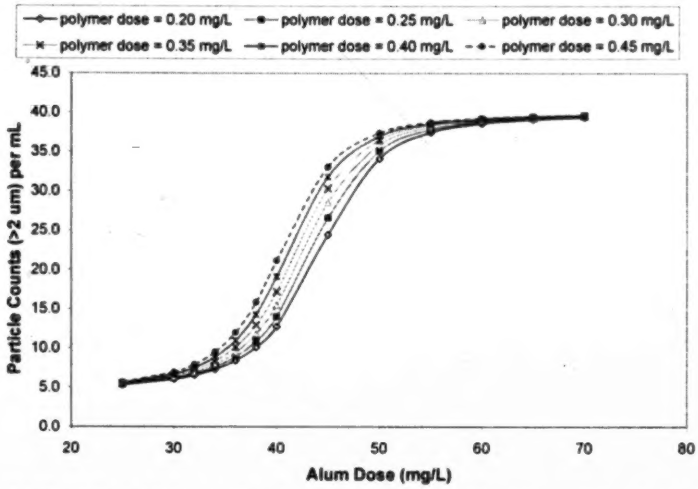


Figure 4-12. The effect of alum and polymer dose on particle counts as predicted by the winter model.

Depicted in Figure 4-12 is the effect of alum and polymer dose as predicted by the winter model. It can be seen that an increase in alum dose can cause a significant increase in particle counts while an increase in polymer dose causes a minor increase in particle counts. Reasons of this behaviour are the same as that described in section 4.2.1.

A much more complex behaviour is observed in Figure 4-13 in which the effect of alum and polymer dose is predicted by the summer model. With polymer doses of 0.12 and 0.14 mg/L, particle counts decrease with increasing alum doses. But with polymer doses of 0.16, 0.20, 0.22, and 0.24 mg/L, particle counts increase with increasing alum dose. Furthermore, the curve for 0.16 mg/L of polymer is different from 0.20, 0.22, and 0.24

mg/L in that a peak is reached with particle counts actually decreasing after an alum dose of 110 mg/L. The reason for this behaviour is unknown, but it is important to note that the overall impact is not significant since particle counts vary only between approximately 4 and 10 counts/mL.

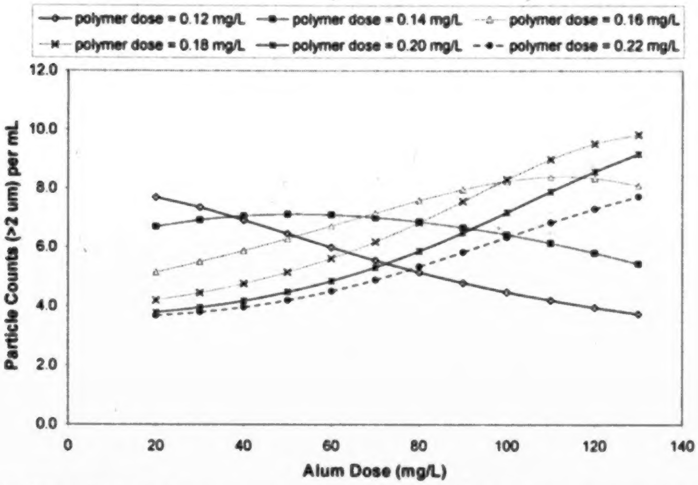


Figure 4-13. The effect of alum and polymer dose on particle counts as predicted by the summer model.

The information presented in Figures 4-11 to 4-13 is another demonstration of the usefulness of utilising the models to assist the operator in determining the treatment action to take. For example, it is significant to note the impact polymer dose has on particle counts compared to alum dose under spring/fall conditions. For instance, if under these conditions a goal of 10 particle counts/mL is desired, a combination of 50 mg/L of alum and 0.28 mg/L of polymer or 25 mg/L of alum and 0.32 mg/L of polymer is required. In opting for the latter, a 50% reduction in alum dose along with only a 14% increase in polymer dose is required as opposed to the former option. Under winter conditions, however, polymer dose is not as significant a factor in finished water particle counts, and care must be taken not to use an alum dose that is too high otherwise, particle counts may be high. Under summer conditions the effect of alum dose and polymer dose is not as significant since particle counts would vary between 4 and 10 counts/mL, therefore the operator could choose a combination of doses that was economically optimal.

Even though the results presented here were limited to operationally controllable parameters, any input used in the models can be examined in this way. This would provide more information to assist operators in determining treatment options. Furthermore, this provides insight into the behaviour of particle counts in response to various parameters, which would be beneficial to researchers in this field.

5.0 Discussion

5.1 Performance Summary

Table 5-1. Statistical performance of the models that were developed.

	spring/fall model	winter model	summer model
R squared	0.79	0.51	0.42
Mean Squared Error	8.9	11.0	5.8
Mean Absolute Error (particle counts/mL)	2.3	2.8	2.0
Maximum Absolute Error (particle counts/mL)	6.7	6.9	5.4

Based on the model performance statistics as summarised in Table 5-1, the spring/fall model performs reasonably well. The winter and summer models have fairly low predicting errors as well, however their R^2 values are not very high. By looking at the model performance graphs in Figures 4-1, 4-2, and 4-3, however, one would have a better idea of how well the models predict. In Figure 4-1, it can be seen that the spring/fall model follows the trends in the data quite well and manages to reach the same peaks, and this is an indication of the model predicting reasonably well. The winter model, however, does not predict as well as the spring/fall model. It can be seen in Figure 4-2, that the model is able to follow the trends in the data but not to the same extent as the spring/fall model. In addition, some peaks are reached such as pattern 3, but others are not such as pattern 6. Altogether this accounts for why the R^2 value is not as high. Nonetheless, it can also be seen that the differences between the actual and predicted particle counts are not that great, and this is reflected in the fairly low mean squared error and mean absolute error. The summer model also does not tend to predict as well as the spring/fall model. In examining Figure 4-3, it is observed that trends in the data are mostly followed except patterns 21-25, and when trends are followed, it is not to the same extent as the spring/fall model. Some peaks are reached such as at pattern 14, but others are missed such as at pattern 4. Since the trends are not being followed as closely and because of some missed peaks and over-predictions, a low R^2 resulted just as for the winter model. However, the actual differences between actual and predicted particle counts are not very high as reflected by the mean squared error and mean absolute error.

5.2 Model Applications

Once the models were developed, a series of examinations into the effects of various factors on particle counts were done using these models and presented in section 4.2. This served to illustrate the great potential these models have as powerful tools that are useful for plant operators and researchers to use. These models are useful in two ways: as a means of determining optimal operational dosages and as a research tool in studying particle counts through the filtration process.

Currently, it is standard practice at a WTP to run jar tests as a means of determining treatment options. Coupled with operator experience, operational dosages are chosen, and this is how process control in the treatment process is achieved. One problem with jar tests is that they do not directly predict filtration performance. Instead, clarification performance of chemicals is measured, which does not necessarily mean that the filters will behave in the same way. Not only are these tests time consuming and inconvenient, but decisions based on experience are not as reliable considering all the variables that are involved that influence filtration performance. Despite the fact that this level of process control generally works for controlling turbidity, particle counts are more variable and harder to control using traditional process control. ANN, on the other hand, can improve process control and has a number of advantages. One major advantage of using the ANN models is that a wet laboratory is not required for performing these tests. Instead, the model acts as a virtual laboratory, and it is very easy to utilise through a spreadsheet program providing treatment options immediately. In addition, there is no concern with problems of scaling up from bench or pilot plant to a full-scale plant because the data that is used to train the models is full-scale data. Figure 4-11 is a strong example of the model's use for determining operational dosages. From the graph, an operator can evaluate and decide the optimal dosage to use. In this example, if an operator wanted a finished water quality of 10 particle counts/mL, the person has the option of using 50 mg/L of alum and 0.28 mg/L of polymer or half as much alum at 25 mg/L and only 14% more polymer at 0.32 mg/L. From this, the operator can decide which option is more economical and yet produces the same result. Process control is achieved quickly and efficiently compared to using jar tests. Furthermore, such determinations are more comprehensive since the impact of multiple parameters can be examined rather than relying on operator experience. A WTP could go one step further and integrate the ANN models as part of the process control system. In a fully automated system, not only does the model give treatment options, but it could also initiate the action providing even more immediate process control action.

Besides providing treatment options, the ANN models are useful as research tools in studying particle counts through the filtration process. Figure 4-4 provides a good example of using the model as a means of researching factors affecting particle counts through filtration. In Figure 4-4, the effect of alum dose on particle counts is examined. In comparing each of the three curves, it can be seen that alum dose affects particle counts differently in different seasons. Researchers can use this information as well as study other factors as demonstrated in Figures 4-5 to 4-13 and make some conclusions as to the extent of effect different factors have on particle counts under different conditions. With an improved understanding of how particle counts are affected, water utilities would be better able to control and reduce particle counts and therefore comply with strict water quality standards.

5.3 Conclusions

It can be concluded that an ANN model of filtration performance can be feasibly developed as shown by the results of the models developed in this study. Given a similar amount of data or more and the same availability of data, the same type of model can be developed for other WTPs.

Another conclusion is that the ANN models developed are useful as tools to assist researchers and WTP operators. As a research tool, the models can be used to study particle counts as demonstrated in this study. It is also useful to WTP operators for determining optimal chemical dosages or to assist in determining courses of action for treatment, thereby improving process control. Not only that, the models that have been developed are unique in that complete plant models are not available. ANN models that have been developed for water treatment focused on unit processes, therefore this was a different modelling approach to encompass an entire plant.

Conclusions can also be made regarding the behaviour of particle counts in response to various conditions. One general conclusion is that particle count behaviour is complex and not necessarily simple and linear. With several factors having an influential effect both individually and combined, it is difficult to make conclusive remarks on overall particle count behaviour. Another general conclusion is that different water conditions result in different particle count behaviour. This can be seen in predictions between the three models, which predict based on data from different seasons. The settings that were used in predictions were those typically occurring within that data set and thus that season. As a result, different curves were observed between each model examining the same factor. For example, in Figure 4-4, the effect of alum dose was examined using each model. The result is three different shaped curves, and it appears to be due to the fact that the set conditions used for each model are different and dependent on season.

More specific conclusions could be made about particle count behaviour as well. Adding non-coagulation/flocculation related chemical to the system, such as PAC and lime, generally increases the amount of particle counts in the finished water. This is illustrated in Figures 4-5 and 4-6. This observation is likely due to the fact that adding these chemicals is equivalent to adding particles to the system, and therefore it is these particles that are contributing to the particle count. The exception appears to be lime dose in the spring and fall where adding more lime decreases particle counts. This may be due to hardness being extensively removed and more particles settling out by adhering to the calcium carbonate precipitate. It could also be concluded that adding excess alum or polymer increases particle counts under winter conditions as illustrated in Figure 4-12. This may be due to particles being added to the system coming from the excess alum and polymer. However, the same behaviour is not observed under spring and fall conditions shown in Figure 4-8 where increased dosages result in decreased particle counts. Particle count behaviour is even more complex under summer conditions and is shown in Figure 4-13. Depending on the polymer dose, excess alum could increase or decrease particle counts. This illustrates the general conclusion that different water conditions result in different particle count behaviour. To be more specific, it appears that temperature and raw turbidity, two major factors that differentiate between the different seasonal conditions, play a crucial role in the way particle counts react to other factors. And this is part of the general conclusion that particle count behaviour is complex because of the multiple factors involved and the possible interactions that occur between the factors.

It is important to stress, however, that the model predictions are limited by the boundaries of the range of data values used to train the models as was previously mentioned in section 4.2. If data that is presented to the models is beyond this range, the resulting predictions are extrapolated and are not necessarily as accurate as opposed to predictions based on data within the trained range. For example, the range of alum dose in the data set for the winter model is between 31-39 mg/L. Thus, in Figure 4-4, the winter predictions outside that range were extrapolated, and particle counts may not necessarily behave in the manner presented in that graph. Therefore, conclusions made based on model predictions outside these boundaries may not be accurate. It would be necessary to study and determine how well the models can predict beyond these boundaries.

5.4 Recommendations

There are a number of issues that would need to be examined in further study. One is the testing of model limits and boundaries as just discussed previously. Another area in association with testing model limits is the verification of model predictions and the conclusions of particle count behaviour made based on these predictions. This would need to be done in the laboratory either bench-scale, pilot-plant scale, or by use of full-scale plants if available. Another issue, particularly with the winter and summer models, is model performance. The winter and summer models do not follow trends as well in the data as the spring/fall model, and this is reflected in their low R^2 values. With the collection of more data providing more data patterns for the models to train from and therefore providing a better representation of the cause-effect relationships in place, model performance should improve especially considering approximately only a year's worth of data is used and split between three models. Another area of improvement would be the Kohonen classification of data. The main problem that arose, mainly in the second stage of categorisation, was the unclear categorisation of some of the data. In other words, the Kohonen classifier could not place certain data patterns into one exclusive category because they possessed characteristics that could have placed them in either category. The determination of the final model configuration is still in question. If an overall model were to be implemented at E.L. Smith WTP, the input data would first pass through a Kohonen classifier and then be fed through to whichever of the three models developed would be appropriate. This is shown in Figure 5-1. However, if the three models were to be combined into one model, the need for a data classifier would be eliminated, and the overall configuration would be much simpler as shown in Figure 5-2.

Associated with that area of study is the need to develop better protocols for merging models since attempts to amalgamate the spring/fall (C1S1) and winter (C1S2) models were unsuccessful.

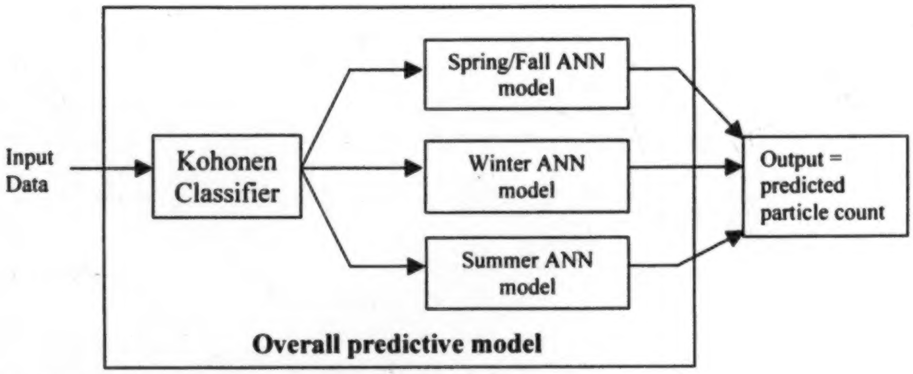

Input Data
Kohonen Classifier
Spring/Fall ANN model
Winter ANN model
Summer ANN model
Output = predicted particle count
Overall predictive model

Figure 5-1. Overall configuration of the predictive model.

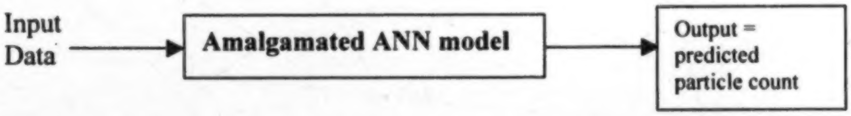

Input Data
Amalgamated ANN model
Output = predicted particle count

Figure 5-2. Configuration of predictive model with all ANN models combined into one.

Once these issues are resolved, the next area of concern would be the integration of the ANN model with the WTP SCADA system. It would need to be determined in what capacity the model would be inserted whether as part of the process control system or as a separate interface apart from the system for the operator to perform "virtual" jar tests.

Although further study is required before full implementation of the models could be done, the study that was done has shown that an ANN model of particle counts through filtration can be feasibly developed. Provided that there is enough data, other WTPs can also develop their own ANN models based on the ones developed in this case. The potential of the model as a useful tool for WTP operators to determine optimal chemical dosages was also demonstrated. Process control would therefore improve, and a water utility could go so far as to incorporating the ANN model as part of the process control system to automate the system thereby improving it. Furthermore, the models developed were shown to be potentially valuable research tools in studying the effects of various factors on particle counts; a field in which much more research needs to be done. With an increased understanding of particle counts, a water utility will be better able to reduce counts and comply with stricter water quality standards.

References

Aim, R. B., S. Vigneswaran, H. Prasanthi and V. Jegatheesan. 1997. Influence of Particle Size and Size Distribution in Granular Bed Filtration and Dynamic Microfiltration. *Water Science and Technology* 36(4): 207-215.

Amirtharajah, A. 1988. Some Theoretical And Conceptual Views of Filtration. *Journal American Water Works Association* 80(12): 36-46.

AWWA. 1990. *Water Quality and Treatment: A Handbook of Community Water Supplies*. New York: McGraw-Hill, Inc.

Boscolo, A., C. Mangiavacchi and F. Drius. 1991. *Artificial Intelligence Tools in Complex Monitoring Systems*. IEEE Instrumentation and Measurement Technology Conference, Atlanta: IEEE.

Castillo, E., J. M. Gutierrez and A. S. Hadi. 1997. *Expert Systems and Probabilistic Network Models*. New York: Springer-Verlag.

Collins, M. R., G. L. Amy and C. W. Bryant. 1987. Evaluation of Factors Affecting Performance of Direct Filtration. *Journal of Environmental Engineering* 113(2): 330-344.

Conlin, J., C. Peel and G. A. Montague. 1997. Modelling of Pressure Drop in Water Treatment. *Artificial Intelligence in Engineering* 11(4): 393-400.

Ebie, K. and Y. Miyake. 1991. Influence of pH Change on Direct Sand Filtration of Kaolinite Clay. *Water Supply* 9: S37-S42.

Evans, J., C. Enoch, M. Johnson and P. Williams. 1998. *Intelligent Based Auto-Coagulation Control Applied to a Water Treatment Works*. IEE International Conference on Control 1998, Swansea, UK: Institution of Electrical Engineers.

Gagnon, C., B. P. A. Grandjean and J. Thibault. 1997. Modelling of Coagulant Dosage in a Water Treatment Plant. *Artificial Intelligence in Engineering* 11(4): 401-404.

Garrett Jr., J. H., D. J. Gunaratnam and N. Ivezic. 1997. Artificial Neural Networks for Civil Engineers: Fundamentals and Applications - Chapter 1. Introduction. *Artificial Neural Networks for Civil Engineers: Fundamentals and Applications*. N. Kartam, I. Flood and J. H. J. Garrett. New York: ASCE.

Ginn Jr., T. M., G. R. Bennett and G. D. Wheatley. 1997. *Particle Counting In Real-World Water Treatment Plant Operations*. AWWA Water Quality Technology Conference, Denver, CO.

Habibian, M. T. and C. R. O'Melia. 1975. Particles, Polymers, and Performance in Filtration. *Journal of Environmental Engineering*: 567-583.

Hall, T. and B. Croll. 1997. Particle Counters as Tools for Managing Cryptosporidium Risk in Water Treatment. *Water Science and Technology* 36(4): 143-149.

Han, T-H., E-S. Nahm, K-B. Woo, C. J. Kim and J-W. Ryu. 1997. *Optimization of Coagulant Dosing Process in Water Purification System*. 36th Annual Society of Instrument and Control Engineers Conference, Tokushima, Japan: SICE.

Hasham, F. A., S. J. Stanley and W. B. Kindzierski. 1998. *Modelling of Urban Air Pollution in the Edmonton Strathcona Industrial Area Using Artificial Neural Networks*. Transportation, Land-Use, and Air Quality: Making the Connection, Portland, OR: ASCE.

Hatukai, S., Y. Ben-Tzur and M. Rebhun. 1997. Particle Counts and Size Distribution in System Design for Removal of Turbidity by Granular Deep Bed Filtration. *Water Science and Technology* 36(4): 225-230.

Jackson, P. 1990. *Introduction to Expert Systems*. Wokingham, England: Addison-Wesley.

Kasabov, N. K. 1996. *Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering*. Cambridge: MIT Press.

LeChevallier, M. W. and W. D. Norton. 1992. Examining Relationships Between Particle Counts and Giardia, Cryptosporidium, and Turbidity. *Journal American Water Works Association* 84(12): 54-60.

LeChevallier, M. W., W. D. Norton and R. G. Lee. 1991. Giardia and Cryptosporidium in Filtered Drinking Water. *Applied and Environmental Microbiology* 57(9): 2617.

Li, S. Y., J. A. Goodgich, J. H. Owens, G. E. Willeke, F. W. Schaefer III and R. M. Clark. 1997. Reliability of Surrogates for Determining Cryptosporidium Removal. *Journal American Water Works Association* 89(5): 90-99.

Lind, C. B. 1997. Particle Counts Can Benefit From Changes in Coagulant Programs. *Water Engineering and Management* 144(5): 24-28.

Liu, J. C. and M. D. Wu. 1997. Fuzzy Control of Coagulation Reaction Through Streaming Current Monitoring. *Water Science and Technology* 36(4): 127-134.

Maher, M-L. and R. Allen. 1987. Expert Systems for Civil Engineers - Chapter 1. Expert System Components. *Expert Systems for Civil Engineers*. M. L. Maher. New York: American Society of Civil Engineers.

Maier, H. R. and G. C. Dandy. 1996. The Use of Artificial Neural Networks for the Prediction of Water Quality Parameters. *Water Resources Research* 32(4): 1013-1022.

Mirsepassi, A., B. Cathers and H. B. Dharmappa. 1995. *Application of Artificial Neural Networks to the Real Time Operation of Water Treatment Plants*. 1995 IEEE International Conference on Neural Networks, Perth, Australia: IEEE.

Moran, D. C., M. C. Moran, R. S. Cushing and D. F. Lawler. 1993. Particle Behaviour in Deep-Bed Filtration: Part 1-Ripening and Breakthrough. *Journal American Water Works Association* 85(12): 69-81.

Muller, B. and J. Reinhardt. 1991. *Neural Networks: An Introduction*. Berlin: Springer-Verlag.

O'Melia, C. R. 1985. Particles, Pretreatment, and Performance in Water Filtration. *Journal of Environmental Engineering* 111(6): 874-890.

O'Melia, C. R. and W. Ali. 1978. The Role of Retained Particles in Deep Bed Filtration. *Progress in Water Technology* 10(5/6): 167-182.

O'Melia, C. R. and W. Stumm. 1967. Theory of Water Filtration. *Journal American Water Works Association* 59(11): 1393-1412.

Patania, N. L., J. G. Jacangelo, L. Cummings, A. Wilczak, K. Riley and J. Oppenheimer. 1995. *Optimization of Filtration for Cyst Removal*. Pasadena, CA: AWWA Research Foundation and American Water Works Association.

Prasanthi, H., S. Vigneswaran and H. B. Dharmappa. 1997. Effect of Particle Concentration on the Entire Cycle of Filtration. *Water Science and Technology* 35(8): 91-102.

Prasanthi, H., S. Vigneswaran, T. D. Waite and R. B. Aim. 1994. Filtration of Submicron Particles: Effect of Ionic Strength and Organic Substances. *Water Science and Technology* 30(9): 149-158.

Rodriguez, M. J. and J. B. Serodes. 1996. Neural Network-Based Modelling of the Adequate Chlorine Dosage for Drinking Water Disinfection. *Canadian Journal of Civil Engineering* 23(3): 621-631.

Rodriquez, M. J., J. B. Serodes and P. A. Cote. 1997. Advance Chlorination Control in Drinking Water Systems Using Artificial Neural Networks. *Water Supply* 15(2): 159-168.

Rumelhart, D. E., J. L. McClelland and PDP-Research Group. 1986. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations*. Cambridge, MA: MIT Press.

Sacluti, F., S. J. Stanley and Q. Zhang. 2000. *Applying Artificial Neural Networks to Predict Cast Iron Water Main Breaks*. Canadian Society for Civil Engineering 2000 Conference, London, Ontario, Canada: CSCE.

Sawyer, C. N., P. L. McCarty and G. F. Parkin. 1994. *Chemistry for Environmental Engineering*. New York: McGraw-Hill, Inc.

Stanley, S. J. and Q. Zhang. 1997. *The Use of Artificial Neural Networks for Process Modeling of Enhanced Coagulation*. Water Quality Technology Conference, Denver, CO: AWWA.

Stark, H. L., S. J. Stanley and I. D. Buchanan. 2000. *The Application of Artificial Neural Networks to Water Demand Modelling*. Canadian Society for Civil Engineering 2000 Conference, London, Ontario, Canada: CSCE.

Stumm, W. and J. J. Morgan. 1996. *Aquatic Chemistry - Chemical Equilibria and Rates in Natural Waters*. New York: John Wiley & Sons, Inc.

Tobiason, J. E., G. S. Johnson and P. K. Westerhoff. 1990. *Particle Size and Filter Performance. Model Studies*. National Conference on Environmental Engineering. Publ by ASCE, New York, NY, USA: ASCE.

Tobiason, J. E., G. S. Johnson, P. K. Westerhoff and B. Vigneswaran. 1993. Particle Size and Chemical Effects on Contact Filtration Performance. *Journal of Environmental Engineering* 119(3): 520-539.

Tobiason, J. E. and C. R. O'Melia. 1988. Physicochemical Aspects of Particle Removal in Depth Filtration. *Journal American Water Works Association* 80(12): 54-64.

Valade, M. T., J. K. Edzwald, J. E. Tobiason, J. Dahlquist, T. Hedberg and T. Amato. 1996. Particle Removal by Flotation and Filtration: Pretreatment Effects. *Journal American Water Works Association* 88(12): 35-47.

Zhang, G., B. E. Patuwo and M. Y. Hu. 1998. Forecasting With Artificial Neural Networks: The State of the Art. *International Journal of Forecasting* 14: 35-62.

Zhang, Q. and S. J. Stanley. 1997. Forecasting Raw-Water Quality Parameters for the North Saskatchewan River by Neural Network Modeling. *Water Research* 31(9): 2340-2350.

Zhu, H., D. W. Smith, H. Zhou and S. J. Stanley. 1996. Improving Removal of Turbidity Causing Materials by Using Polymers as a Filter Aid. *Water Research* 30(1): 103-114.

Zhu, X. X. and A. R. Simpson. 1996. Expert Systems For Water Treatment Plant Operation. *Journal of Environmental Engineering* 122(9): 822-829.

Appendix

Hach Model 2100N Laboratory Turbidimeter

Principle of Operation

This instrument is a Nephelometer with the capability to measure with either the Ratio on or Ratio off. It meets the design criteria of the United States Environmental Protection Agency and is acceptable for compliance reporting. The optical system is comprised of a tungsten-filament lamp, lenses and apertures to focus the light, a 90 degree detector, forward-scatter light detector, and a transmitted-light detector. Turbidity measurements at less than 40 NTU are performed utilising only the 90 degree scattered-light detector or using the complete set of detectors (ratio). With the Ratio ON (necessary for samples greater than 40 NTUs), the instrument's microprocessors uses a mathematical calculation to ratio signals from each detector.

Specifications

Principle of Operation: Nephelometric

Measurement Units: Nephelometric Turbidity Units (NTU), Nephelos, European Brewery Convention (EBC)

Ranges (With Ratio ON)
NTU Mode: 0-4000 NTU with automatic decimal point placement or 0-0.999, 0-9.99, 0-99.9, and 0-4000 with manual range selection
Nephelo Mode: 0-26 800 with automatic decimal point placement or 0-9.99, 0-99.9 and 0-26 800 with manual range selection
EBC Mode: 0-980 with automatic decimal point placement or 0-0.999, 0-9.99, 0-99.9, and 0-980 with manual range selection

Ranges (With Ratio OFF)
NTU Mode: 0-40
Nephelo Mode: 0-268
EBC Mode: 0-9.8

Accuracy: ±2% of reading plus stray light from 0-1000 NTU; ±5% of reading from 1000 to 4000 NTU based on Formazin primary standards and with Ratio ON, ±2% of reading plus stray light from 0-40 with Ratio OFF
E.L. Smith Acceptable Limits: ±20% at < 10 NTU and ±10% at > 10 NTU
Reference Conditions: 0 to 40°C, 0 to 90% RH Noncondensing @ 25°C, 115/230 Vac ±17%, 50/60 Hz

Resolution: 0.001 on lowest range
Repeatability: ±1% of reading or ± 0.01 NTU, whichever is greater

Reference Conditions: 0 to 40°C, 0 to 90% RH Noncondensing @ 25°C, 115/230 Vac ±17%, 50/60 Hz

Response Time: 6.8 s with signal averaging off or 14 s with signal averaging on

Standardisation: Formazin Primary Standards

Display: 5-character LED, 13.7 mm high digits with custom annunciators

Light Source: Tungsten filament lamp. Lamp life 8800 h (typical)

Signal Averaging: Operator selectable on or off

Sample Cells: 95 mm high x 25 mm diam. Borosilicate glass with rubber-lined screw caps.

Sample Required: 30 mL minimum

Secondary Standards: Gelex Secondary Standards

Temperature
Storage Temperature: -40 to 60°C
Operating Temperature: 0 to 40°C
Sample Temperature: 0 to 95°C

Operating Humidity Range: 0 to 90% RH Noncondensing @ 25°C; 0 to 75% RH Noncondensing @ 40°C

Instrument Stabilisation Time: 30 min with ratio on, 60 min with ratio off, typical applications leaves instrument on 24 h/d.

Measurement Procedure

1. Collect a representative sample in clean container. Fill sample cell to the line (Approx. 30 mL) and cap the sample cell.
2. Hold sample cell by the cap and wipe to remove water spots and fingerprints.
3. Apply a thin bead of silicone oil from the top to the bottom of the cell enough to coat cell with a thin layer of oil. Use provided oiling cloth to spread oil uniformly and wipe off excess.
4. Place the sample cell in the instrument cell compartment and close the lid.
5. Select manual or automatic ranging by pressing the Range key.

6. Select the appropriate signal averaging setting (on or off) by pressing the Signal Avg key.

7. Select the appropriate Ratio setting (on or off) by pressing the Ratio key

8. Select the appropriate measurement unit (NTU, EBC, or NEPH) by pressing the Units key.

9. Read and record the results.

Met One Model PCX Particle Counter (with local display)

Principle of Operation

Particle count measurement is online and continuous. Water is directed into the sensor and funnelled through an optical flow cell measuring 750 x 750 microns. Flow rate through the sensor is at 100 mL/min. Each particle that passes through the sensor generates a signal corresponding to its size. Each sensor comes with a calibration curve showing the signal response versus size of each sensor. NIST-traceable spheres of known size are used to calibrate each sensor. Information is stored in the memory of the sensor and is used to separate the particle counts into the proper size category.

Specifications

Smallest Particles Counted: down to 2 microns

Largest Particles Counted: up to 750 microns

Distance from Computer to Sensor: 4000' maximum (for entire RS-485 signal path)

Power: 115 VAC (±10%); Optional 220 VAC (±10%; 50/60 Hz)

Enclosure: NEMA 4X-Rated

Indicators: Power, Particle/Alarm, Calibration Status, Count Display

Flow Rate: 100 mL/min, nominal

Max. Pressure: 65 psi, not more than 1 min duration; 55 psi continuous

Fluid Connections:
Inlet: Quick disconnect. Connects to ¼ inch OD Tubing

Outlet: Quick disconnect. Connects to ¼ inch OD Tubing

Accuracy:
Manufacturer: ± 5 counts / mL for a sample of blank water
E.L. Smith: ± 10 counts /mL for filtered water, ± 10% for raw water

Other Parameter Measurements

Raw Water Temperature

An online thermometer measures the temperature at an accuracy of ±2%.

Total Hardness and Total Alkalinity

Readings are based on a colorimetric titration by a Tytronics FPA 400 series analyser with an accuracy of ±5 mg/L

The following steps outline a typical analysis cycle:

1. Fluid from the sample stream is used to wash out the reaction cell by exchanging several volumes of fluid
2. The sample is captured using a syphon method that assures volume repeatability better than 1%
3. The appropriate reagent(s) is added and the solution is mixed
4. Titrant is slowly added, and the time to reach an end point is measured
5. The sample concentration is calculated and results displayed

pH

Readings are measured by a Rosemount Model 1054 online pH analyser, which utilises a glass electrode. The manufacturer's stated accuracy is ±0.01 pH units, however E.L. Smith's acceptable limit is ±0.2 pH units.

Alum Dose and Polymer Dose

Alum and polymer feed is measured with a magmeter (magnetic flow meter) with an accuracy of ±0.5%.

Lime Dose and PAC Dose

Lime and PAC feed is measured with a massmeter (mass flow meter) with an accuracy of ±0.25 to 0.5%

Raw Flow

Measured using a magmeter (magnetic flow meter) with an accuracy of 0.5% except at low flows in which accuracy slightly decreases.

Calculation of R^2

The formula that NeuroShell 2 uses for calculating R^2 is:

$$R^2 = 1 - \frac{SSE}{SS_{yy}}$$

where

$$SSE = \sum (y - \hat{y})^2$$

$$SS_{yy} = \sum (y - \bar{y})^2$$

y is the actual value (particle counts in this case),

$\hat{y}$ is the predicted value of y,

$\bar{y}$ is the mean of the y values

R^2 is defined as the coefficient of multiple determination. It compares the accuracy of the model to the accuracy of a trivial benchmark model wherein the prediction is just the mean of all of the samples. A perfect fit would result in an R^2 value of 1, a very good fit near 1, and a very poor fit less than 0. If model predictions are worse than one could predict by just using the mean of the sample case outputs, the R^2 value will be less than 0.

Definition of Mean Squared Error, Mean Absolute Error, Maximum Absolute Error

Mean squared error is defined by NeuroShell 2 as a statistical measure of the differences between the values of the outputs in the data set and the output values the network is predicting. This is calculated as the mean of the square of the actual values minus the predicted values, i.e., the mean of (actual - predicted)^2. The errors are squared to penalise the larger errors and to cancel the effect of the positive and negative values of the differences.

Mean absolute error is the mean of the series of absolute errors calculated, which is defined as absolute value of the difference between the value the network is predicting for an output and the actual value of that output. In other words, it is the mean of |actual - predicted|.

Maximum absolute error is the maximum of the series of absolute errors calculated.

The lower or nearer to 0 these errors are, the more accurate the model predictions.